DOCUMENT RESUME

ED 243 709 SE 044 480

AUTHOR	Beattie, Ian D.; And Others
TITLE	Research Reports from the Seventh National Conference on Diagnostic and Prescriptive Mathematics. 1981 Research Monograph.
INSTITUTION	Research Council for Diagnostic and Prescriptive Mathematics, Kent, OH.
REPORT NO	ISBN-0-940466-07-4
PUB DATE	Mar 82
NOTE	212p.; Papers presented to the RCDPM Annual Conference on Diagnostic and Prescriptive Mathematics (7th, Vancouver, Canada, April 1980).
AVAILABLE FROM	Research Council for Diagnostic and Prescriptive Mathematics, 441 Beryl Drive, Kent, OH 44240.
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	Algebra; Computation; *Diagnostic Teaching; Educational Research; Error Patterns; *Learning; *Learning Problems; *Mathematics Education; Prediction; Preservice Teacher Education; *Remedial Mathematics
IDENTIFIERS	*Mathematics Education Research

ABSTRACT

This document contains 14 papers selected from those presented at the Seventh Annual Conference of the Research Council for Diagnostic and Prescriptive Mathematics. Categorized as interdisciplinary studies are reports on "Mathematics as an Activity of Our Intelligence" by Skemp and a study on "The Brain Basis of Developing Mathematical Skills: Application to Learning Disabilities" by Kinsbourne. Studies related to young children include one paper on predicting achievement from first-grade interviews by Callahan and Boller and a review of research on young children's interpretation of mathematics pictures by Campbell. Studies of testing and validation include papers by Englehardt, Van Wagenen, and Thomas; McGregor; and Beattie and Harley. Studies on errors and understanding are reported by Rees, Hart, Crawford, DeVincenzo, and Bright and Harvey. Young presents a study on the effects of instruction and field experience with preservice elementary teachers. Finally, Rising discusses "Diagnosis and Prescription for the Research Council on Diagnostic and Prescriptive Mathematics." (MNS)

****
*   Reproductions supplied by EDRS are the best that can be made   *
*   from the original document.   *
****

ED243709

RCDPM

1981
RESEARCH MONOGRAPH

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
● Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Research Reports from
The Seventh National Conference
On Diagnostic and Prescriptive Mathematics

EDITED BY

Ian D. Beattie, Tom [illegible],
James [illegible] III and Douglas [illegible]

UNIVERSITY OF VANCOUVER

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Carol Larson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

RESEARCH COUNCIL
FOR
DIAGNOSTIC [illegible]

SE044480

COPYRIGHT, 1982, BY THE RESEARCH COUNCIL FOR DIAGNOSTIC AND
PRESCRIPTIVE MATHEMATICS
FIRST PRINTING, MARCH, 1982, AT BOWLING GREEN, OHIO

PRINTED IN THE UNITED STATES OF AMERICA

ISBN-0-940466-07-4

ALL RIGHTS RESERVED. NO PART OF THIS MONOGRAPH MAY BE REPRODUCED OR TRANSMITTED IN ANY FORM OR BY ANY MEANS, ELECTRONIC OR MECHANICAL, INCLUDING PHOTOCOPY, RECORDING, OR ANY INFORMATION STORAGE AND RETRIEVAL SYSTEM, WITHOUT PERMISSION IN WRITING FROM THE PUBLICATION CHAIRPERSON.

PUBLICATIONS OF RCDPM ARE PRODUCED UNDER THE AUSPICES OF THE PUBLICATIONS COMMITTEE OF THE ORGANIZATION. THE 1980-81 COMMITTEE MEMBERS WERE: DON BERNARD, CHAIRPERSON; GRACE BURTON; EARL LEGGETTE; CAROL LARSON; AND JAMES BRUNI.

THE MATERIALS PRINTED IN THIS MONOGRAPH ARE THE EXPRESSION OF THE WRITERS AND NOT NECESSARILY A STATEMENT OF POLICY OF THE RESEARCH COUNCIL FOR DIAGNOSTIC AND PRESCRIPTIVE MATHEMATICS.

ADDITIONAL COPIES OF THIS OR ANY PUBLICATION OF RCDPM MAY BE PURCHASED FROM RCDPM HEADQUARTERS, 441 BERYL DRIVE, KENT, OHIO 44240.

THE RESEARCH COUNCIL FOR

DIAGNOSTIC AND PRESCRIPTIVE MATHEMATICS

1981 MONOGRAPH

SELECTED PROCEEDINGS AND RESEARCH REPORTS

of the
SEVENTH ANNUAL CONFERENCE
Vancouver, April 1980

EDITED BY

IAN D. BEATTIE

TOM BATES

JAMES SHERRILL

AND

DOUGLAS OWENS

UNIVERSITY OF VANCOUVER

BRITISH COLUMBIA

Preface

Membership reaction to the Vancouver conference was very favourable. Whilst the Vancouver setting helped set the mood, undoubtedly the basis for that judgement was the excellence of the program. This excellence reflected some brave decisions by the conference committee, such as continuing a trend from previous conferences to include speakers from other, but impinging disciplines, and inviting the best possible speakers even though travel expenses would be high. The result was a varied program, interesting speakers and excellent participation.

The papers included in this volume were chosen not only for their excellence but to convey the flavour of that conference, not an easy task. With that goal in mind and the constraints of space some papers could not be included. To their authors I extend my regrets for the omissions. The editing proved to be far more time consuming and difficult than anticipated, and I wish to thank those of my colleagues who so willingly assisted. Tom Bates edited the Rising paper, James Sherrill edited the Engelhardt and McGregor papers and Douglas Owens edited the Callahan and Campbell papers. Each also edited papers which did not make it into this volume.

It was impossible to edit these papers without being mindful of those parts of the Rising paper regarding publications and editing. I suppose that means that one of his goals is already achieved. Not all members may agree with his views or even with the decision to include his paper, but from our organization's inception there has been discussion among members regarding its direction and even its name. I feel the net result of the paper will be a strengthening of the R.C.D.P.M.

I wish to acknowledge, also, the marvellous cooperation of our departmental secretary Beth Buchanan who prepared this manuscript.

IDB

Contents

Preface i

INTER-DISCIPLINARY STUDIES

Mathematics as an Activity of our Intelligence
Richard R. Skemp 1

The Brain Basis of Developing Mathematical Skills: Application to Learning Disabilities
Marcel Kinsbourne 13

STUDIES RELATED TO YOUNG CHILDREN

Predicting Mathematics Achievement from Entering First Grade Interviews
Leroy G. Callahan and V. Christine Boller 24

Review of Research: Young Children's Interpretation of Mathematics Pictures
Patricia F. Campbell 39

STUDIES OF TESTING AND VALIDATION

Locus of Control and Feedback in Diagnostic Mathematics Testing
Jon M. Englehardt, R. Keith Van Wagenen and Noel Thomas 52

Review of Research on the Validation of Learning Hierarchies
John D. McGregor 68

The Development of a Diagnostic Test of Subtraction of Whole Numbers
Ian D. Beattie and Dwight D. Harley 87

STUDIES OF ERRORS AND UNDERSTANDING

An Exploratory Study of Mathematical Abilities: The Learning and Teaching of Some Basic Mathematical Difficulties
Ruth M. Rees 100

Mistakes in Mathematics - The Research of CSMS
Kathleen Hart 111

Difficulties and Errors in Computation - A Report and Commentary on Some Recent Studies in England, Canada and the U.S.A.
Douglas H. Crawford 118

The Relation Between Elementary Algebra Students' Errors in Arithmetic and Algebra in Selected Types of Problems
Mary Ann DeVincenzo 133

Diagnosing Understandings of Equivalences
George W. Bright and John G. Harvey 144

TEACHER EDUCATION

The Effects of Instruction and Field Experience in Diagnostic-Prescriptive Teaching of Mathematics on Preservice Elementary Teachers
Sharon L. Young 170

INTROSPECTION

Diagnosis and Prescription for the Research Council on Diagnostic and Prescriptive Mathematics
Gerald R. Rising 182

Mathematics as an Activity of our Intelligence
A Model for Diagnosis and Remediation of Learning Difficulties in Mathematics

Richard R. Skemp
University of Warwick

From the early 1960's onwards, I was intuitively convinced that any intelligent child, or adult, could learn mathematics if he was willing to work at it, and was properly taught. Surely the main ability required for mathematics is the ability to form and manipulate abstract ideas? And surely this ability coincides closely with what we mean by intelligence? This view was supported by factor analytic studies showing the highest correlations between mathematical ability and 'g'. There did, however, seem to be an excessive number of exceptions, of intelligent children and adults who seemed unable to grasp mathematics, however hard they tried. So the alternative view, that mathematics was a specialised ability which could be almost entirely lacking in people who were otherwise quite intelligent, had to be taken seriously. I did not agree with the latter view, but neither was I in a position to refute it.

Since problems of learning and teaching are psychological problems, this led to a long period of work in the psychology of learning mathematics, of developing these ideas into a theory of intelligent learning, and of teaching which would help this to take place. For it became ever more clear that mathematics was not the only subject that was badly taught and ill understood. It just showed up more clearly in mathematics.

What was most lacking in that early work was a satisfactory relationship between the learning of concepts and schemas on the one hand and motivation and action on the other. The field of cybernetics supplied the answer. Goal directed action began to appear both as a foundation concept and as a unifying concept around which many others began to fall into place.

There is a close parallel here with the new significance which the concept of set has taken on in mathematics within our own memories. It is used both as a foundation concept and, with the help of others which are closely related, as a unifying concept by which whole areas of mathematics, previously a collection of separate topics, are now integrated into a coherent whole. My own first encounter with that concept was in a first year university course. Now children encounter it right at the beginning of their mathematics. I and, I suspect, my teachers at that time would have been surprised if someone had told us of the ways in which that concept would be put to use.

For the years 1973-1978 I was engaged on a similar exercise in a different area of thinking, that of intelligent learning, using goal-directed activity as a foundation and a unifying concept. About two thirds way through I began to see in general outline how mathematics fitted into this new approach as an important special case. And that is mainly what I am working on now.

Using hindsight I can see why it took me so long. I started from the wrong place. The conception of intelligence which I learned as a student, and which has dominated psychology for seventy years, based on I.Q. and its measurement, has remained isolated from theories of learning. Likewise, the theories of learning which I learned to pass my examinations failed to take account of the factor of intelligence. Only in recent years have I come to realise that intelligence is a kind of learning. It is the highest form of learning, in which man most differs from other animals. It's not something of which people have a certain amount, and the problem is how to measure it. It is something people do and the problem is to find out how it functions.

Again using hindsight, if the learning of mathematics is closely related to intelligence, and if intelligence describes an activity, just as memory describes the combined activities of receiving, storing, and retrieving information, then it follows that unless we understand the function of intelligence - what it does and how it does it, we shall not be in a strong position to diagnose or remedy students' problems in the learning of mathematics. If we want to reach an understanding of how mathematics is learned, and how we can teach it better, the place we have to start from is a model of the functioning of human intelligence.

To offer such a model is an ambitious undertaking, which represents a right-angle change of direction to that of the majority of psychologists in this field. Whether you accept it or not, I hope that you will find it of interest, and stimulating to your own thinking. I shall outline the model as briefly as I can, emphasizing the parts most relevant to children's difficulties in learning mathematics. In order to do this I shall have to leave out aspects which I would like to include, such as the role of emotions. The model is a synthesis of ideas from various sources and some new ones of my own.

The starting point is the assumption that much human activity is goal directed. I do not claim that all our behaviour is goal directed, but that some, perhaps much, is, and that to ignore that fact is to ignore one of the most important features of behaviour. If we want adequately to understand what people are doing we need to go beyond the outward, easily observable, aspect of their actions and ask what their goal is. To say that Richard Skemp is riding his bicycle may be perfectly true, but it is only part of what matters. I may be riding to my office; or I may be enjoying physical recreation with my son; or I may be testing the adjustment of my three-spead gear. Moreover, going to my office may underlie actions which look quite different. For example, instead of bicycling I might be standing still by the side of the road, waiting for a bus; or I might be walking, which in this case would also be along a different path; or I might be driving my car, in which case I would start off in a totally different direction. To limit a description of what was happening to the observable behaviours, superficially very different, would be to miss what they had in common, namely the goal state. And if we can identify a goal underlying certain observable behaviour, it is reasonable to look for the significance of that goal rather than some other.

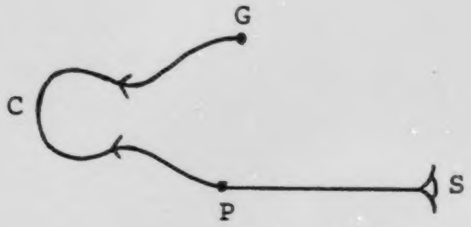
Let us assume then, that our actions are not random but are systematically directed towards bringing about goal states. Now, the situations in which we find ourselves doing this are not always the same. Nevertheless, we are able to achieve our goals by varying our actions appropriately. A system for doing this I call a director system ; and since the idea is borrowed from cybernetics, the ways in which a director system works will probably be somewhat familiar. But I need to recap them briefly for the sake of what comes after.

Its essence is a comparison between the present state of some operand and its goal state, combined with a plan of action directed always so as to reduce this difference until the present state coincides with the goal state. This means that both the present state and the goal state have to be represented within the director system in some way: otherwise, how can they be compared?

We need a sensor,(S) which takes in information about the present state (P) of the operand, and represents it within the system

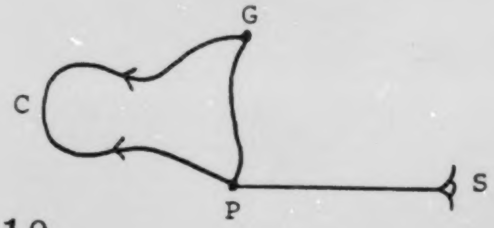
We need a representation of the goal state, (G)

And a comparator, (C) which compares these.



While changing from its present state to the goal state, the operand passes through a number of intermediate states. Each of these must be represented within the director system in order for the process of comparison to continue, without which the director system cannot continue successfully to direct action.

This is equivalent to saying that not only the initial state and the goal state, but also a path from the present state to the goal state, must be represented within the system.

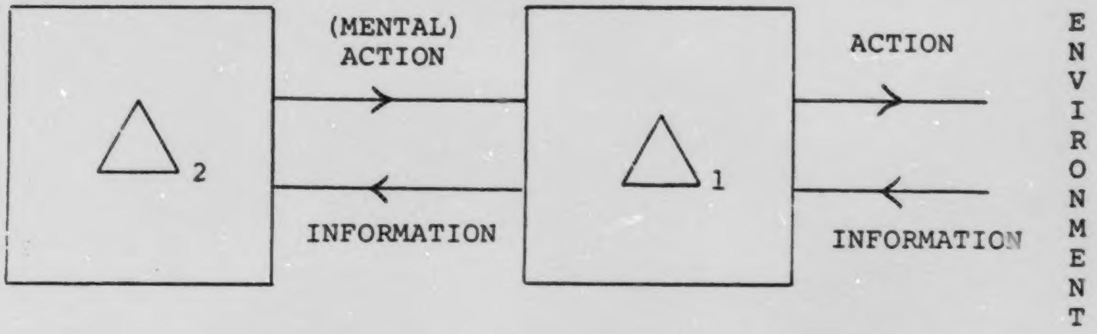


In the lower animals, many of these director systems are innate. They form part of the survival equipment of the species. But there is an upper limit to what can be transmitted genetically; and there are other disadvantages, such as slowness to adapt to environmental changes. So it is not surprising that some species have evolved the ability to set up new director systems during the lifetime of an individual, and to improve the ones they have. This is how I now conceptualise learning: as a change in an organism's director system towards a state of better functioning. Other animals can learn too, of course. But we have so far outdistanced other animals in our ability to learn that I regard our own species as an evolutionary breakthrough. It is not so much that we can do better than other animals at the same kind of learning; at maze learning, for example, rats are better than humans. But we have available a more advanced kind of learning, which is qualitatively different from the kind exemplified by maze learning. It's the ability to learn in this more advanced kind of way that I would now call intelligence.

Intelligence is a kind of learning which results in the ability to achieve goal states in a wide variety of conditions, and by a wide variety of paths. And if we've been able to agree that action, much of it, is not random but goal-directed, you may be willing to take with me the further step of conceiving learning as goal-directed also, or at least some kinds of learning.

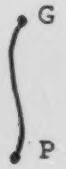
The new model uses the concept of a director system of two levels.

Delta-one (Δ_1) is a director system whose operands are physical objects in the environment.

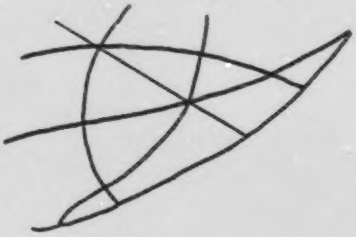


Delta-two (Δ_2) is a second order director system which has delta-one as its operand. Its function is to take delta-one to states in which delta-one can do its job better. It does this in several ways.

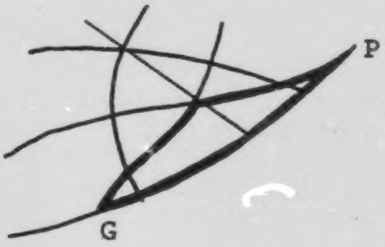
We have already seen that there has to be, represented within delta-one, a path from present state to goal state.



A simple path of this kind may be innate, or it may be learnt. Either way, it is closely tied to action; and a particular sequence of actions at that. There's not much adaptability; and if one gets off the path, one doesn't know what to do to get back on. So it's much better to have a cognitive map, , within which can be represented a variety of present states and goal states.



and from which a great number of possible paths can be derived, from any initial state to any goal state, provided that both can be represented within the same cognitive map.



It is learning at this level with which intelligence is particularly concerned, and although its function is one step removed from action, it makes action more likely to succeed because it is better adapted for each particular task.

Intelligence thus contributes to adaptability in two ways:

1. by the construction of these cognitive maps - not just one, but a large number, for different kinds of jobs that delta-one does, and

2. by deriving from these particular plans appropriate to different initial states and goal states. These plans can then form the basis of goal-directed action as already outlined.

Next, I would like to look at the nature of these cognitive maps in greater detail. We can think of them as mental models of certain features of the outside world, without which goal-directed activities cannot take place.

SCHEMA CONSTRUCTION

SCHEMA BUILDING	SCHEMA TESTING
Mode 1	Mode 1
from our own encounters with actuality: experience.	against expectations of events in actuality: experiment.
Mode 2	Mode 2
from the schemas of others: communication.	comparison with the schemas of others: discussion.
Mode 3	Mode 3
from within, by formation of higher order concepts: by extrapolation, imagination, intuition: creativity.	comparison with one's own existing knowledge and beliefs: internal consistency.

The left-hand column corresponds to changes of state of our schemas; the right-hand column corresponds to activities of the comparator in delta-two, testing that these changes are for the better. The three modes of building, and the three modes of testing, can be used in any combination; and we usually do use more than one of them at a time.

The possible uses of our schemas also fall into three categories.

USES TO WHICH SCHEMAS CAN BE PUT

Group 1.	For directing action by delta-one.
Group 2.	For making possible co-operation.
Group 3.	As agents of their own growth.

As with the categories of building and testing, these may occur singly or in various combinations. Connections can also be found between the three modes of construction, and the three groups of uses. One of the areas in which the present model helps our thinking is the ways in which all of these activities may interact.

Now, as Heraclitus has told us, 'We cannot step twice into the same river'. The present experiences from which, sometimes, we learn become part of our past, and will never again be encountered in exactly the same form. But the situations in which we need to apply what we have learned lie in the future as it becomes present, or as by anticipation we bring it to our present thinking. It follows that if our mental models are to be of any use to us, they must represent, not singletons from among the infinite variety of actual events, but common properties of past experiences which we are able to recognise on future occasions. A mental representation of these common properties is how, for many years now, I have described a concept; and for this process of concept formation I use the term 'abstraction'. Concepts represent, not isolated experiences, but regularities abstracted from these. It is only because, and to the extent that, our environment is orderly and not capricious that learning of any kind is possible. A major feature of intelligent learning is the discovery of these regularities, and the organising of them into conceptual structures which are themselves orderly. These conceptual structures, or schemas, are like cognitive maps only more so. We could think of them as cognitive atlases, in which a large dot representing, say, Vancouver on a map of Canada can itself be expanded into another map, a street map of the city, in which this hotel now appears as a dot. But we could expand this dot, if we liked, into a three-dimensional plan of the hotel. So a schema is a cognitive map in which each point has interiority. Schemas of this kind provide our delta-one systems with mental models which are useable for many different situations, and from which a variety of plans can be derived as required. They are essential for the functioning of delta-one.

Constructing these schemas for use by delta-one is the job of delta-two. How delta-two sets about this particular goal-directed activity is what I want to look at next.

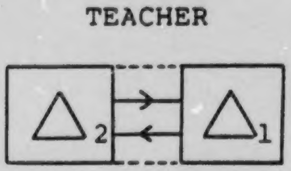
Constructing includes both building and testing. If one is building a wall, a delta-one activity, building consists of putting on mortar and bricks, and testing consists of checking their position for spacing, alignment, and verticality. The ways in which we build and test our schemas, a delta-two activity, fall into three main categories, as indicated below.

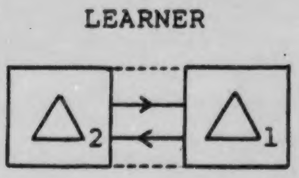
Application to Diagnosis and Remediation

Working out the applications of the new model looks like being a long job. In the time remaining, I shall talk in general terms and keep more detailed aspects for other occasions.

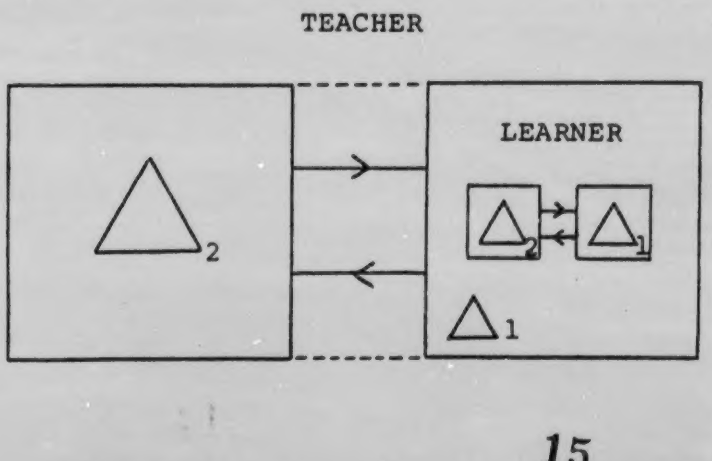
A first major emphasis of the model is that behaviour is not shaped by the environment, in the sense of being a product of the environment. People adapt their behaviour to achieve their goals, either the same goals, or equivalent goals, in spite of changes in the environment. So when A intervenes in the goal-directed activity of B, we have an interaction between two director systems which is not only more complicated, but qualitatively different, from the kind of interaction implied by a stimulus-response model. One of the areas to which the model directs our attention is: 'what are the conditions for intelligent co-operation between two goal-seeking persons?'

Co-operation in a learner-teacher situation is complicated by the fact that not two, but four director systems are involved.





The learner's delta-one/delta-two form a closely coupled interactive pair, by which learning has been going on since birth. No one taught us to speak! Like digestion, another invisible process, a child's built-in learning system works very well on its own for much of the time. Intervention is as likely to do harm as good unless we have a clear mental model of how it works, whether we are intervening in bodily or mental processes. So, a teacher must have, as a necessary part of his professional expertise, a model of the learner's delta-one/delta-two system.



This global diagram conceals great interiority. I'll expand it a little, since it was only when I began to analyse a teacher's task in terms of this that I began to realise the difficulty and complexity of a teacher's professional task.

(i) First, he must have this model available at a high level of generality, representing the regularities of intelligent learning common to different topics and different learners. This will include a model of the various ways by which delta-two acts on delta-one to take delta-one from a present state to a goal state.

(ii) Within this, he must be able to realise any present state of any learner's schemas, so that he can

(iii) devise plans by which he can co-operate with the learner's delta-two to take the learner's delta-one from its present state to the next goal state.

The teaçher knows what this learning goal is like much better than the learner - he has been there already. He should also know a suitable path, sequence of states, in this case states of the learner's schema, to this learning goal from the learner's present state. He should also know the ways of schema construction, building and testing, to bring into action, or rather, to try to arrange for the learner to bring into action, to take the learner's delta-one along this path.

Co-operating with the learner's delta-two involves giving it a chance to act. That means giving it time to reflect, supporting its own goal-directed activity, rather than continually bombarding it with information and demands. It is the learner's delta-two that has to do the work. Often it does a good job on its own. Good teaching is helping it to do its work, and this includes sometimes standing back and not intervening unless asked.

If all this sounds difficult, it is. It is more difficult than I have made it sound, not less. The more I examine the nature of a teacher's profession, the more I am coming to think that it is one of the most difficult there is to do really well. Under-estimating these difficulties may be part of our present problem. It is also hard to escape the conclusion that it may sometimes be better not to teach someone at all than to teach them badly.

Additional difficulties in the teaching of mathematics result from the nature of the subject itself. Its difficulties are closely related to its strengths. I shall mention two of these briefly, and spend more time on the third.

(i) Mathematical schemas are of a high order of abstraction.

Strength. They provide very adaptable, multi-purpose models which can be used by delta-one to direct actions in many different universes of discourse. E.g. the same basic mathematical model a/b = c , with appropriate units attached, can represent either the relationship between distance, time, and speed, or that between mass, volume, and density, or it can represent Ohm's law.

Difficulty. Mathematical concepts are remote from primary concepts, being derived from secondary concepts which are themselves derived from secondary concepts, etc. In other words, mathematical concepts don't embody regularities of our sensory experiences of the physical environment, but regularities of these regularities, and relationships between them, at a high order of abstraction.

Suggested remedy. By working in the opposite direction, analysing the higher order concepts which are learning goals, we can identify what lower order concepts are pre-requisite for each new higher-order concept. These lower order concepts need to be analysed in their turn, right down to primary concepts. We can then draw a dependency network, and use this to guide the sequencing of our teaching.

(ii) Mathematical schemas are of high information content, all of the concepts having much interiority.

eg. The algebraic statement $a(x+y) = ax + ay$
contains an infinity of statements like $5(7+11) = 5\times7+5\times11$

Strength. This kind of organisation maximises the amount of information we can store, and minimises cognitive strain by allowing us to retrieve as little as is necessary for the direction of action at a given time. This is one of the reasons why mathematical models are so powerful throughout science and technology, industry and commerce.

Difficulty. It is very easy to overlook the high information content of mathematical statements just because they say so much in so little. The result is that mathematics teachers often give out information faster than it can be taken in. My own teachers at university certainly did this.

Suggested remedy. Realise this. Cover the ground more slowly, and allow more time for assimilation and consolidation. Teach less and teach it better.

(iii) Mathematics makes no use of mode 1 building and testing, once we have left the early stages.

Strength. This eliminates a great deal of noise, especially that due to experimental error. It concentrates on manipulation of the mental model, which, unrestricted by practical difficulties, can be developed to a high degree of expertise. The latter may eventually lead to major payoff in mode 1. That is, the use of mathematical models can enable us to achieve goals with physical operands which would otherwise be impossible. However, success or failure in this is not the way we normally test the mathematics itself. What we test in the science laboratory is a particular theory or model, formulated in mathematical concepts.

Difficulty. The learning, communicating, and using of mathematics is thus completely dependent on symbolic processes. These are indispensible as combined labels for distinguishing concepts, and handles for manipulating them. But in combination with the fact that interaction between teachers and learners also relies heavily on mode 2, we have a major source of potential difficulty which pervades mathematical instruction. This I should like to discuss more fully as a final point.

Mathematics, as an example of intelligent learning, is essentially a delta-two activity of two kinds:

(i) the building up of mathematical schemas, and

(ii) deriving particular plans from these.

But teacher-learner interaction is a delta-one activity: albeit of a rather special kind, namely the utterance of symbols. By 'utterance' I mean speaking, writing, putting on an overhead projector, etc. This opens wide the door to a serious misconception on the part of learners about the nature of their learning goals. They are all the more likely to go through this door, and spend their time thereafter in a mental realm which they think is mathematics but isn't, because of the dual nature of their teacher's authority towards them.

We need to distinguish two kinds of authority. Though the categories are quite general, they are particularly relevant to the teacher-learner relationship. These are positional authority, and sapiential authority. Positional authority is the kind held by a policeman, an army officer, a works manager. He gives instructions which others obey. It is essentially a power relationship, more overt in some cases, more politely disguised in others. Sapiential authority is authority of knowledge, which other people consult to help them to achieve their own goals. Examples of sapiential authorities are medical doctors, lawyers, technical experts. These are in a co-operative relationship with their clients. The clients consult the expert's knowledge, but are then free to choose what use they make of it.

A school teacher's authority is of both these kinds. To create and maintain the orderly conditions necessary for learning to take place, he exercises positional authority. But in helping children to learn mathematics, his authority is that of knowledge. The children's position is, or should be, that of clients, though I do not expect everyone to agree with this view.

This raises an important and general question, which belongs to a much wider universe of discourse than our present one: to what extent are these two roles of a school teacher compatible? I shall not follow up on this point, but there is another consequence which is central to our present concerns.

From a subordinate towards someone in positional authority, the appropriate question is: 'What do you want me to do? How am I to behave?' So children, who are subordinate to a teacher in positional authority, are likely to perceive their task predominantly in terms of learning to behave the way teacher wants them to. Some of this behaviour consists of asking permission before leaving the room, raising a hand before speaking. Some of it is behaviour of a more specialised sort, namely the speaking or writing of certain mathematical symbols. This way, the learning of mathematics becomes an exercise in obedience, of an increasingly complicated kind.

We arrive at the same possibility from a different beginning. Since interaction between learner and teacher is by means of symbols, the door is open to misconception on the part of learners that the subject they are meant to be learning is concerned with just that - speaking or writing the correct symbols: correct, as judged by the teacher. To the extent that a child perceives the teacher's authority as positional, he is likely mentally to go through that door. Some teachers may even given them a push, by treating a mathematical error as if it were a misdemeanour. But I concur with the view of a senior inspector for mathematics that taking a mathematical error should attract no censure, whereas misbehaving so that the learning situation is interfered with should quite rightly attract censure.

A child who gets stuck with this conception of mathematics, as the memorisation of rules without reasons for the manipulation of symbols without meanings, is in for a hard time: because ultimately the memory load becomes impossible. In terms of the present model, he is two steps away from learning mathematics as an activity of his intelligence, by which he would find the task much easier and more enjoyable. Did I say two steps away? I think it is more like two universes.

Reference.

Skemp, Richard R. Intelligence, Learning and Action: A Foundation for Theory and Practice in Education. Chichester: John Wiley and Sons, 1979.

The Brain Basis of Developing Mathematical Skills: Application to Learning Disabilities

Marcel Kinsbourne, M.D.
University of Toronto

In discussing learning disabilities we must consider the broadening consequences of what initially can be quite a limited selective unreadiness for a particular form of instruction. The child who enters first grade unready in some important respects can fail. This failure leads to a lowered self image for the child, disappointment on the part of the parents and frustration on the part of the teacher. While the fuss is all about the child in question, siblings feel ignored and class peers feel slighted. These factors further lower the child's self image and that, plus the failure, make for underachievement, deepening into a totally suspended effort. The consequences of this cycle of events are negative for both intellectual and personality development. Social promotion deepens the failure with a final result of illiteracy or other type of incapability and underemployment that does not do justice to the child's native abilities. The pressure on the child, and his resentment of that pressure lead to progressive alienation which typically manifests itself in one of three styles, acting out of anger, withdrawal from the situation or a denial that anything is happening. These are short term coping devices, but because the problem is long term, these devices become institutionalized into personality characteristics and the individual ultimately adopts a maladapted life style, perhaps isolated or delinquent. So a quite limited failure within an otherwise intact intellectual repertoire may, because of what society expects and the pressures put on those who do not meet these expectations, cause very serious problems in living in the long term.

So when we discuss learning disabilities we are simply discussing the child who is hard to teach for intrinsic reasons, not for environmental reasons, not because of emotional pressures like impaired family dynamics, but because he or she, although he has most mental operations available at the expected level for the age, does not have some of them available and those that are not available happen to be crucial for some academic area. This is really a case of misfit and the area of learning disability is the area of individual misfit. Had the child available all the mental operations appropriate for the age, that would be a normal child. Had the child none of them available, that would be a retarded child. These children who are hard to teach are all different in what they don't understand, what they don't concentrate on, what doesn't interest them and what they can't remember. In a group setting it is virtually impossible to instruct these children effectively. So another way of defining learning disability, the most pragmatic way, is to say that these are children whose needs for instruction are individualized. Nothing will be achieved unless there is individualized confrontation between teacher and pupil. The problem is an educational problem and after discussion of other matters, I shall return to the educational setting.

Going back into the history of learning disabilities one can easily detect a dissonance which still exists. There is at this time, as you well know, some opposition between some educators and some learning disabilities specialists. To simplify the positions, educators say these children are simply hard to teach and the solution is to give them better teachers. Learning disabilities specialists say that since the educators are not doing it, another way must be found. And that is what happened eighty or so years ago when physicians entered the act. They saw children who were not learning as expected for their general intelligence, said these children had a disease and the disease was that they could not read. There was a reason for that point of view. It happened against a background of an explosion of knowledge in adult neuropsychology, and in the previous few decades a variety of syndromes had become apparent. When a person is damaged in some limited area of the brain, he or she loses only certain mental skills manifested in certain limited incapabilities. From these facts people began to build up a geography of the brain, and there in among the geography was so-called alexia, damage to the centre that subserves reading. This view, where the brain is a depository of faculties, we call faculty psychology and we do not quite see it this way anymore. But, once you admit of an acquired alexia, presumably something an adult has and loses on account of damage, a child might fail to develop in the first place. So it seemed reasonable to look among children for the non-development of these very same capabilities which, when they do develop, are vulnerable to brain damage subsequently. And a child who seemed intelligent and could not read was diagnosed to be a child who could not read, a case of congenital alexia.

This circularity of reasoning has been a major problem in the area of learning disabilities all along. What tends to happen is that a child who has trouble trying to read or calculate sees one or an assorted committee of experts and having been evaluated, (because in North America we don't just examine, test or diagnose, we evaluate) the very value of the person is evaluated to be dyslexic or dyscalculic, as well as he can't read or calculate, and brain damaged to boot. This approach is based on a bewilderment that I find bewildering. People are surprised, to the extent that they can barely believe, that one can be both smart and dumb at the same time, that one may have certain capabilities more developed at a given age or in adulthood than others. They act as if they truly believed that intelligence was all of a piece, that you could take humans such as those in this room and arrange them in one dimension, from the cleverest to the least clever. What an extraordinary thing to suppose! Humans are multi-dimensional, but multi-dimensionality of physical characteristics is thought not to apply to the mental case. Yet mental development is just as irregular as physical develoment. There is nothing surprising about somebody being verbally gifted and spatially inept or spatially skilled and mathematically not gifted and so on. These combinations obviously arise from genetic diversity, but if one does not recognize that this is so, one has to look for other explanations, and there has been an intense look for other explanations.

There have been several agendas here and it is both interesting and tricky to put out the threads of argument. There is the agenda that, if it can be determined why a person though otherwise clever is in some respects dumb, a disease will have been diagnosed and, by implication, a disease mechanism can be uncovered. Then, perhaps after suitable investigations of the mechanism, the brain malfunction can be corrected. That is a medical model. You look for something that can be changed. Alternatively you can depart from the surface manifestations such as the problem in the academic area and delve into the underlying problem, which is presumably neurological or perhaps neurochemical, something to do with functional neurons, because that is what controls behaviour. Correct the neurons and the learning will look after itself. The correction can be either curative or symptomatic where the manipulation must be maintained. In either case the learning is better. This is also a medical model. Anybody who has pneumonia is grateful to the medical model, but this may not be the same type of case.

A second reason for looking for a specifically identifiable physical neurological syndrome is more pragmatic and a little bit more subtle. It tries to escape from the circularity of reasoning alluded to earlier which is not satisfying at all. In this agenda, if you can identify some constellation of deficits that characterize those who cannot read or calculate for developmental reasons, then you have shown something over and above restating the complaint. So what is looked for is a syndrome or significant associates. A syndrome would be not only the complaint but several things also wrong which can be explained by some underlying deficit of mental operations. Or you look for associated abnormalities such as signs of brain damage, the idea being that, if there are other signs of brain damage, then the thesis that the child's learning problems are due to brain damage would be thereby made more acceptable. There are two approaches here. The first looks for minimal brain damage or brain dysfunction, M.B.D. There have been many studies of children who are selectively poor learners by some definition, and there are a variety of definitions as you know. Children may be compared with normal or excellent learners with regard to a variety of non-cognitive measures. Examination may show neurological abnormalities, electro-encephalographic abnormalities or changes in body shape or development. In fact, it has been pretty reliably found that whatever you are looking for is there. If, for example, you do a multiple congenital anomalies count, poor learners have more of them than good learners. But there are poor learners who are physically immaculate and there are those who have all the stigmata, but are good learners. So you have lots of Type I and Type II errors. Or you can look at another kind of 'defective' population, left-handers, and find that, in whatever subject, there is a higher proportion of left-handers among the poor learners than there is in the general population of good learners. None the less, more poor learners are right-handed than left-handed. A second approach is the search for soft neurological signs. A soft neurological sign is the persistence of a reaction or appearance which would be normal in a younger child but should have disappeared by the age of the child in question, in other words signs of neuro-motor immaturity. The argument is that, if the neurological control of motor behaviour is immature, it can more readily be believed that a learning problem also is due to an immaturity elsewhere in the brain, namely that part which controls learning. This again is a guilt- by-association argument, for though it is often true that these children do have more soft signs than do controls, it is subject to many Type I and II errors. This in my mind makes this approach sterile at least in terms of practical spinoff. After all, there are children with cerebral palsy who have, by all accounts, massive neurological problems who learn well and vice versa.

So, perhaps a more intelligent search is the search for typical constellations of deficits that are intelligible on a brain-based rationale but which wouldn't be expected to arise in their full specificity on account of emotional blocks, anti-intellectual atmosphere of the home, bad relationship with a particular teacher, and so on when you would expect more generalized difficulties. So, we begin to look for specific patterns, and we base the nature of our search, because one always has to look for something over nothing, on what we learn from the adult. So, adult neuropsychology then gives you a number of model syndromes that are potentially present in our children. We take our target population, the poor learners, and see whether we can find developmental analogs of those adult syndromes in the children. I am going to mention a number of these syndromes, but I would first remind you that these are intelligible in terms of what we know about how the brain is organized in very general terms, namely the differential contributions of two hemispheres, left and right. And this is what we know.

The left hemisphere is specialized for sequential, specific indication or attribution. In other words, the left hemisphere makes distinctions. It is in control of the momentary focus of attention or mental awareness, one thing at a time, in sequence. The right hemisphere is entrusted with relational information, how the specifics picked up by the left hemisphere relate to each other. As I look around this room, although I have a general view of its size, dimension, the density of people, illumination, and so on, if I decided to close my eyes and tried to specify how somebody looked or what he was wearing I would find myself at a loss to say more than one or two words. This is because detailed focal attention is sluggish and moves like a pinhole from point to point. So as I look around, I am really looking around with a pinhole but I don't have that phenomenological impression. That is because the right hemisphere enters each successive focus into a framework, a stable framework, of how they each relate to the others. So whereas the left hemisphere gives me successive information, the right hemisphere gives me a simultaneous panorama against which that successive information is entered. Now, there are two conclusions that we can draw from these propositions if they are accepted as correct. One is that the two hemispheres make disparate contributions to the same overall effort. In other words they separately elaborate parts of experience which then become coordinated. The left hemisphere by itself would give you free floating features, (FFF's), a nose here, a slightly balding head there, whereas the right hemisphere would give you a framework without content. I can not even begin to imagine what it would look like. So the two hemispheres do not represent alternate ways of doing the same thing. The second conclusion, that they do represent alternate ways of doing the same thing, is what one might call exploitation psychology. You must have heard the claims that there are two consciousness inside the same head, that whatever you are doing you might do better if you took a different hemispheric style and so on. It does not make much sense to say that if you were more right hemispheric, if you noticed spatial relations more, but not what relates to what, you would be better off. The reason that the specializations are apart in the head is precisely that they have to get together. If the job is to choose alternatives, there is no need to separate their brain bases because they won't interfere, you are only choosing one. Whichever one you choose, the other one is not chosen, you can learn that program. The difficulty arises when you are doing two things at the same time within a network, and the brain is a highly connected network. It's not a computer at all. It's a network. If you want to do two things at the same time in a network which are based on different program characteristics, the big difficulty is crosstalk, interference, and that's when you have to have them on different sides of the head so you can put in an inhibitory barrier conveniently, insulating each from the other till they are properly elaborated and then coordinated. So hemisphere specialization doesn't offer you alternative life styles. On the contrary, it offers you a coordinated single life style just as your experience is unitary, although it is built up of these separate components.

So, one point I want to make then is that with respect to how one should teach, how one should think, how one should teach thinking, we have absolutely nothing whatever to learn from hemispheric specialization. We knew it all before; we knew that people can be logical, analytic. We knew they could be spatially synthetic, we knew they could be creative, we knew they could be rigid, and we knew they could be accurate and precise or fuzzy. The fact that someone claims that when you are being fuzzy you are being fuzzy in a particular place does not make you more or less fuzzy. It is not useful information and methodologies cannot be based on that kind of information in my view. A second point of interest that arises is that we can now understand the various syndromes, examples of which I will give because the brain works in this way. It has dual operation types that it can perform but in different parts of it, that it applies to different problems. The sequential facility in the back of the left hemisphere is applied to specifying input stimuli percepts. In the front it is applied to programming a sequence of outputs. It is still sequential but sequential to different ends. So, if you have damage to one part you might lose the sequential capability for one kind of operation, while damage to another part of the same hemisphere would lose the sequential capability for another kind of operation. In the right hemisphere, we have this relational facility, but if it is damaged you do not necessarily lose relation skills for everything. Damage to one part may lead to impairment in map reading, to another, impairment in putting together a sketch or a diagram, and to a third, impairment in recognizing a person's face which is an act based largely, if not entirely on spacial relationships between features. So you have different syndromes depending on the locus of the damage.

What kinds of syndromes, then, have actually been observed that might relate in some way to mathematics? Actually knowledge in this area is not nearly as good as that in the areas of speaking and reading, writing and understanding speech, but some things are known. First of all, an obvious point. If one has aphasia, a difficulty in speaking and understanding speech, this obviously will affect mathematical instruction. The explanations will not be intelligible and the person will have trouble saying or writing, reading or understanding the number sequences or the algebraic symbols used in the currency of mathematics instruction. It turns out that actually naming digits and identifying digits is the easiest thing for any human who is aphasic to do. The digit names are, for some reason, so overlearned and so well known that whatever deficit exists in using words, they are the least affected. They are still affected, but the least affected. Whatever deficit exists in reading, reading the digits is the least affectd but it's affected. Now, take the example I mentioned before of alexia, difficulty in reading words, difficulty in reading numbers, in reading complex numbers. To say somebody has alexia is saying something rather sterile, making an unjustifiable assumption about how the brain is organized, namely, that there is a part of the brain that deals with reading as a reading centre. Now, that seems unintelligible because reading after all is the coordinated outcome of multiple potentially identifiable operations and there's no reason to presuppose that they all take place in the same part of the head. Why should they? On the contrary, they are likely to take place in different parts of the head because they have to happen together for the reason I have just explained. Each operation must have its own personal space to occur in prior to coordination. So, in fact, to say that it's alexia doesn't do justice to the phenomenology that are there are other alexias, that there are different reasons why adults might lose ability to read, including the reading of numbers. After brain damage, for instance, there are those adults who can look at the symbol, perceive it but have lost its reference. They can not retrieve. They have lost the mental habit of interpreting the symbol in a particular way. There are others who in fact can do that, but are very much slowed down in doing that. What that does is to break up the fluency of the reading, such that they can only read a letter at a time or a digit at a time. This makes reading difficult when there are multiple letters on digits involved in the total production, for, as they go along they forget the ones they started with. The word is no longer a word, the unitization of the symbols is destroyed and each letter or digit is its own unit. So many people with left hemispheric damage are easily overloaded with multiple units, and it would appear to be reasonable to look in children who have problems with use of symbols for the analog of this kind of deficit. And I can simply say it has not been found, though I give it as an example of what has been investigated.

Another syndrome you can look at is that of error patterns, the argument being that the defective performance will give a clue as to which component they no longer can bring to that performance. This can be done in two basic ways. You can either look at errors of commission such as what they do when they should do something else or you can look at which manner of instruction they find easiest. As an example of the second approach, children with reading problems have been classified into what are called dysphonetics and dyseidetics. Dysphonetics make phonetic mistakes, like-sounding mistakes. Dyseidetics make visual confusions in reading and writing, misreading letters for similar ones. You can either conclude that the dysphonetics have trouble using phonics, or you can conclude that they have trouble using whole word approaches and therefore are thrown back on using phonics even where it is difficult, which is why you get phonic mistakes. You can conclude that the dyseidetics have trouble with the visual images or you can conclude that, because they cannot use the phonic approach, they overload the visual stores and that is why you get those visual mistakes. In other words, since you do not know their teaching history, how well they were taught, which approach and at what time, there is not a single thing you can conclude. So you have to forget about error patterns.

Individual errors may be more revealing. It has been shown that adults brain damaged in different places make two kinds of spelling mistakes. Some patients made a high proportion of extraneous letter mistakes, putting letters into words when they were not in the words but sounded as if they should be. Others tended rather to get letters in the wrong serial order. This seemed to have some potential applicability to children for several reasons. The people who made the extraneous letter errors were the aphasics, and among children there are those who have trouble learning to read, who clearly have a language problem. These are children whose development was slow in the first place, and is somewhat imperfect even when they read aloud, although that may have been missed because it is subtle.

Then there is another group, those with the Gerstmann Syndrome, who do not seem to have this problem with the choice of the correct speech sound or its graphic representation, but have a problem with getting it in the right order on paper. They are interesting because they also have a problem with calculation. Now, of course the aphasics have a problem with story problems, but that is obvious. The Gerstman Syndrome has four elements, finger agnosia, right-left orientation, dysgraphia and dyscalculia, but there is a common problem in computing underneath these four rather different sounding kinds of difficulties, a problem with identification based on relative position, on sequential position. In arithmetic the problem becomes one of place value and children have trouble with the 56, 65 type of reversal and carrying from column to column. The problem is not in number sense, the problem is in place value. This syndrome is a good illustration of how you can show some validity in making an organic attribution by showing a constellation of surface manifestations. The specificity of the syndrome gives you some assurance that it is brain based.

Another type of difficulty that impinges on mathematics has to do with geometry, with the representation of shapes. In adults there is a very interesting difficulty called constructional apraxia, which is simply a difficulty in copying drawings, and geometric figures are usually used. The thing that is interesting is that you can get constructional apraxia either after left damage or right damage, but the nature of the problem differs. If asked to copy the representation of a cube, in a clear cut case of left damage, an individual will draw three squares or a big square partitioned into three surfaces. The figure is simplified and everything becomes rectangular with the acute and obtuse angles regressed to the right angle. It is what I call rectangular vision and this is also normal for four year olds when given the same task. Those who have trouble for right-sided reasons make quite a different kind of error pattern. They draw the components correctly, a square, a parallelogram, another one, but they misarticulate them grotesquely. It is like an exploded representation of parts that are not relationally integrated. This is as you might expect from what the right hemisphere does. With children I have never seen that kind of right hemisphere deficient behaviour, suggesting that in child development the right hemisphere is a little bit ahead of the left in that respect so there is slack there to be taken up. What limits children is the left side of things. They behave like left-damaged adults either when you give them a copying task which is harder than their age warrants, if they are normal, or when the task is one they should be able to do but they are not normal. In the latter case, consider a four year old who can draw right angles, but can not do acute angles even though he can draw the component straight line and oblique line separately. You can try this with your four year old, the limitation when it is developmental is that salient. And I have not yet discovered how to teach that. I am not saying it cannot be done, but I do not know how. Just doing it over and over is not it and it is not a motor problem. It is something of a developmental thing. They deal in right angles. Mentally the programme is rectangular, boxes in the head and that is restated by certain forms of brain damage in adults. And the older child with a deficit is like a younger normal child with respect to that particular operation.

What does all that tell us about the difficulty, since we knew ahead of time that the child could not draw acute angles? Well, it tells us that if the mental operation is immature, the only way to deal with it is to provide for the child the mental structure that the child does not provide for himself. You cannot teach him to provide the structure, you cannot teach him the mental operation though you can teach him tricks which is sometimes a very useful practical device. You can drill, you can do rote teaching, but it will not generalize. That is the problem. But in order to know what mental structure to provide, you have to know the child's mental development, and I feel that what is most lacking in the education of teachers is good mental development instruction, perceptual development, insightful aquaintance with the way in which abilities grow, what children can do and why, at the different ages. Unless you can put yourself in the child's head and can look at the world in that way you make dreadful mistakes.

Take, for example, a typical Piagetian non-conservation example with two parallel aligned lines. The child sees them as being the same length, but when one is moved he now says that one is longer. As I understand it, Piaget explains this by saying that the child does not yet have a sense of the principle that moving something can not change how long it is. I do not believe that for a moment. I am sure the child is giving an answer based on present information. The previous information does not come into the decision because young children are overly dominated by existing context. Right through developmental psychology there is evidence of this everywhere. They act under stimulus control. What happened a minute earlier was not misunderstood; it just isn't represented now. They are under immediate control. You can see this in sentence comprehension. You can say to a child, "The dog chased the cat." And he will understand it, and then you can say, "Cat chased dog the." And he will understand it. He understands it because the dog is a stimulus, the cat is a stimulus, chase is a stimulus, and the child ignores the word order, the syntax and gets the message anyhow. There are many illustrations of this. That is why what happened a moment ago does not change the judgement that this line is longer. It is longer because it looks longer, not to you, but to the child. You cannot imagine how things look to the child, because what is different between you and the child is how you attend to things. And there is no way of attending other than how you yourself are attending or imagining, for imagining is in itself an act of attention, of imaginary attention. The thing you can least understand is the deficit of attention, how things look other than they look, how they can look other than they look. The reason that children make this mistake is that they attend only to one end of the display. Since one line sticks out farther it is judged longer. How could you, even if you tried, attend to only one end of the display? This teaches you that when you show children something, you do not know what they are seeing. You have to find out what they are attending to in the display. It will not be the same as what you are attending to.

Good illustrations of this come from reversals of letters or digits. Some rather dumb things have been said about this tendency such as, those who make reversals have inverted or reversed vision. But those who make reversals do not go out the window when they want to go out the door. Moreover, if you are born with inverted vision you have normal vision. This is what I call the El Greco fallacy. It has been said that because El Greco drew long thin people he was astigmatic, that that was how he saw them. But there is a fallacy here, of course. Since he is matching what he is seeing and what he is producing, if they are both distorted they cancel out. So El Greco was not astigmatic. The reason that children normally reverse, and that some children reverse even to an age of nine or ten when the normal age when reversals occur is five or six, has to do with what is called the perceptual hierarchy, with what they notice about a display. Any stimulus, however simple, has multiple attributes. The common ones are shape, size, location, colour, orientation and sequential position. They are all present simultaneously, all perceptible to the child, but he does not attend to them all at the same time. Those up on the hierarchy preempt attention. He may notice and therefore remember the colour or configuration, but not the other attributes and will make reversal errors. Note that the orientation is not misperceived, it is just not taken into account. It is an attention limitation, not in time as in the previous non-conservation example, but in regard to multiple attributes of a figure. And the behaviour with regard to orientation is random, sometimes a reversal is made, sometimes not, with roughly equal probability. If you properly orchestrate the situation, these children are capable of noting the orientation and reproducing it correctly. If you remove those attributes which are more salient, the latent ability to deal with orientation appears. The child's attention is directed towards a desired attribute not by pointing out, not by underlining, but by removing distractors. In this, of course, you are fighting an unequal battle against the makers of textbooks. Without distractors all textbooks would look alike.

So the principles of teaching are very simple and they are ones which teachers know. First, when you are dealing with a child about whom you can make no prior assumptions, not a normal child, not a retarded child but a child who is individually variable, everything has to be stated in full detail. The one explanation that can not be misunderstood is a complete explanation, one that states every component. The second point is that you must never overload, never give a second item of information unless you are sure the first has been understood, and you must have constant feedback to check on the understanding. The third point is that you can have no constraints on the time taken, the bell is irrelevant. You may stop before the bell if the child is tired or go through it if some point is being learned. The fourth point, one that I have written about more than the previous ones, is that learning is not enough. Overlearning is essential. You are teaching a skill and fluency is the essence of any skill. The skill should not just be available in terms of being correct: the response should be at minimal latency and reliable across all contexts. Unless you overtrain on step one, step two will interfere with step one. When something has been learned, in part the organism has changed, with a different response propensity. Until that point, comes a moment of anxiety, an alien environment, the learning falls off like the emperor's clothes. So you need overlearning of the first point, then the second and so on in graduated fashion. This sounds like rote drilling which can be done in two ways. Psychologists who study verbal learning use rote drilling to overload the organism and so deliberately produce a sample of errors to study. Psychologists use drill to get people to do things wrong. Teachers, on the other hand, use drill so that people will get things correct, so that they will have a readily available skill. This means teachers should drill in a cumulative fashion, a, ba, cba, and so on, always back to make sure that a is still there to do c.

So, we have come full circle to fairly close to where we should, having taken an excursion through neurology, neuropsychology and child development. We come back to the fact that a slow learner needs a slow teacher. If an explanation is not understood we need a better explanation. In other words, we need excellent teaching and excellent teaching is teaching adapted to the specific areas in which the child has difficulty. Everything else is tangential.

Predicting Mathematics Achievement from Entering First Grade Interviews

Leroy G. Callahan
V. Christine Boller
State University of New York at Buffalo

Over the past three years students from a large urban school district were interviewed individually on entry into first grade. Interview questions elicited student performance on number and measurement tasks. The primary reason for the interviews was to assist first grade teachers identify mathematical strengths and weaknesses of their students during the first few weeks of the school year. As a by-product of that endeavor, responses of each student were retained and coded in order to explore factors from the interviews that might be valid predictors of future mathematics achievement.

The students were a cross-section of entering first-graders from a large urban school system. The racial mix was approximately half black and half white. There was a similar mix of boys and girls. The first group of 950 entering first graders was interviewed at the beginning of the 1976-77 academic year. The number of interviews increased in subsequent years. Approximately 4800 interviews were carried out over a three year period.

In the spring of 1979 scores on standardized achievement tests in mathematics from first through third grades were collected on approximately 400 students of the 950 students interviewed during the 1976-77 school year. These data were used in ascertaining whether performance on the interview tasks on entry to first grade predicted subsequent achievement as measured by standardized mathematics achievement tests. The following research reports on some initial results from exploring the prediction question.

In considering the prediction of achievement Lavin (1965) pointed out that before an empirical relationship can be viewed as established with a high degree of reliability, repeated observations are necessary and rather consistent findings must emerge. The performance of three successive groups of entering first graders on the interview tasks presented the opportunity to study the stability of performance for large numbers of students. Tables 1 through 13 present descriptive information on students' responses to interview questions. There are three sets of repeated observations of students entering first grade, 1976-78. Included in the tables are results of readministration of the interviews at the end of the respective school years for a sample of about 140 students each of the three years. These post-test data from the three years were combined and reflect a total of about 420 interviews of students leaving first grade. On the following pages the interview questions and tabled summaries of responses are presented.

Task #1. "Start at one and show me how far you can count."

	N	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
1976-77	950	3%	17%	25%	20%	7%	4%	1%	1%	1%	20%
1977-78	2234	3%	18%	25%	17%	7%	4%	2%	1%	1%	20%
1978-79	1633	2%	14%	25%	21%	6%	5%	2%	2%	1%	22%
Post Test	420	1%	2%	8%	6%	4%	3%	2%	2%	1%	70%

TABLE 1: Point of Termination in rote count from one

Task #2. "Can you start at five and count backward (down)?" (Same for 10 and 13)

	N	From 5	From 10	From 13
1976-77	950	66%	59%	25%
1977-78	2234	63%	57%	26%
1978-79	1633	63%	57%	25%
Post Test	420%	91%	91%	66%

TABLE 2: Ability to Count Back from 5, 10, 13

Task #3. "Show me 4 blocks." (Same for 8, 10, 13 blocks)

	N	4	8	10	13
1976-77	950	96%	90%	86%	75%
1977-78	2234	96%	88%	84%	72%
1978-79	1633	97%	88%	83%	72%
Post Test	420	100%	98%	98%	94%

TABLE 3: Ability to Demonstrate
4, 8, 10, 13 blocks

Task #4. When shown a trio of dot cards with N-1, N, N+1 (random order) dot cards the direction was, "Show me the card with 3 dots." (Same for 6, 10, 13)

	N	3	6	10	13
1976-77	950	96%	88%	67%	56%
1977-78	2234	96%	87%	66%	58%
1978-79	1633	96%	86%	64%	58%
Post Test	420	100%	97%	83%	77%

TABLE 4: Ability to discriminate 3, 6, 10 dot cards

Task #5: Given a deck of 12 cards, 4 with 3 dots, 4 with 6 dots, 4 with 10 dots, all in different arrays, and shown 1 card for each number, the question was asked, "Can you place the remaining cards in the right piles?"

3 DOTS	1976-77	950	3%	2%	3%	92%
	1977-78	2234	2%	2%	2%	93%
	1978-79	1633	2%	1%	4%	93%
	Post Test	420	0%	0%	1%	99%
6 DOTS	1976-77	950	4%	6%	13%	73%
	1977-78	2234	4%	5%	11%	80%
	1978-79	1633	4%	6%	11%	79%
	Post Test	420	0%	0%	4%	96%
10 DOTS	1976-77	950	6%	7%	14%	73%
	1977-78	2234	5%	6%	13%	76%
	1978-79	1633	5%	6%	15%	74%
	Post Test	420	0%	0%	4%	96%

TABLE 5: Ability to recognize 3, 6 and 10 through various arrays

Task #6: (A) When given a set of seven dot cards (arrays of 1-7 dots) the direction given was, "Can you put these in order from littlest to biggest starting with the one card?"

(B) When given a set of seven numeral cards (1-7) the direction given was, "Can you put these in order from littlest to biggest starting with the one card?"

	N	Dot Arrays 1-7	Numeral Cards 1-7
1976-77	950	72%	83%
1977-78	2234	80%	84%
1978-79	1633	75%	85%
Post Test	420	95%	95%

TABLE 6: Ability to order seven array and numeral cards

Task #7: Given a paper with numbers printed on it, the direction given was, "Read these numbers for me." (1, 2, 3, 6, 9, 7, 10, 15, 18)

	N	1	2	3	6	7	9	10	15	18
1976-77	950	98%	98%	98%	87%	87%	79%	86%	51%	57%
1977-78	2234	99%	98%	98%	90%	89%	80%	88%	54%	63%
1978-79	1633	99%	98%	98%	90%	89%	82%	88%	55%	64%
Post Test	420	100%	99%	99%	99%	99%	99%	99%	95%	95%

TABLE 7: Ability to read numbers presented visually

Task #8: With a paper and pencil on the table, the student was given the direction, "Take the pencil and write the numbers I give you on the paper (3, 8, 4, 9, 10, 11, 15, 24)

		N	3	4	6	8	9	10	11	15	24
CORRECT	1976-77	950	77%	80%	67%	80%	58%	74%	70%	55%	24%
	1977-78	2234	75%	80%	69%	81%	58%	75%	73%	53%	27%
	1978-79	1633	70%	79%	70%	82%	59%	76%	74%	49%	28%
REVERSAL	1976-77	950	17%	11%	14%	--	19%	--	--	2%	6%
	1977-78	2234	20%	10%	13%	--	20%	--	--	2%	8%
	1978-79	1633	23%	10%	13%	--	18%	--	--	4%	7%
TRANSPOSED	1976-77	950	--	--	--	--	--	11%	--	5%	2%
	1977-78	2234	--	--	--	--	--	11%	--	5%	3%
	1978-79	1633	--	--	--	--	--	10%	--	4%	3%
UNABLE OR WRONG NUMBER	1976-77	950	6%	9%	18%	20%	22%	14%	30%	38%	66%
	1977-78	2234	5%	10%	17%	19%	21%	14%	27%	40%	62%
	1978-79	1633	7%	11%	17%	18%	23%	14%	26%	43%	62%
	Post	420	99%	99%	97%	98%	97%	98%	96%	96%	90%

TABLE 8: Ability to write numbers from oral presentation

Task #9: "What is this called?" (Pointing to each shape)

	N	Square	Triangle	Rectangle	Circle
1976-77	950	85%	85%	62%	92%
1977-78	2234	84%	83%	62%	92%
1978-79	1633	84%	81%	61%	92%
Post Test	420	93%	89%	80%	98%

TABLE 9: Ability to identify and name geometric shapes.

Task #10: "Which coin is a penny?" (Same for nickel, dime and quarter). "Which would buy the most?" "Which would buy the least?"

	N	Penny	Nickel	Dime	Quarter	"Most"	"Least"
1976-66	950	95%	73%	82%	75%	90%	45%
1977-78	2234	95%	69%	82%	72%	89%	47%
1978-79	1633	96%	68%	82%	74%	88%	48%
Post Test	420	99%	86%	93%	88%	95%	75%

TABLE 10: Ability to identify coins and distinguish "most" and " least" quantities

Task #11: (Hand child a ruler) "What is this called? What do you use it for?"
(Hand child one 3" ruler and the 12" ruler) "Which is longer? Shorter?"

	N	ID	Measure	Draw	Hit Point	Longer	Shorter
1976-77	950	81%	64%	7%	10%	99%	98%
1977-78	2234	81%	59%	10%	1%	98%	97%
1978-79	1633	79%	78%	6%	9%	92%	99%
Post Test	420	94%	80%	82%	-	100%	100%

TABLE 11: Ability to identify and name ruler and its function
Ability to discriminate "long" and "short"

Task #12: (Child is given 2 clay balls) "Which one is heavier or are they the same?"
"Which one is lighter?"
"How much do you weigh? How would you find out?"

	N	Heavier	Lighter	No. Res.	0-25 lbs.	26-75 lbs.	76-100 lbs.	100 lbs.	Function
1976-77	950	76%	76%	51%	10%	37%	2%	-	49%
1977-78	2234	72%	73%	55%	10%	33%	1%	-	48%
1978-79	1633	78%	82%	32%	5%	58%	3%	1%	68%
Post Test	420	88%	83%	36%	8%	53%	3%	-	75%

TABLE 12: Ability with heavy/light: Personal Wt., function of a scale.

Task #13: (Hand child clock face) "What is this called? What is it used for?" (For hands at 8 and 5 o'clock) "What time is this?"

	N	ID Clock	Function	8:00	5:00
1976-77	950	98%	95%	45%	46%
1977-78	2234	98%	94%	45%	48%
1978-79	1633	99%	94%	51%	42%
Post Test	420	99%	98%	87%	85%

TABLE 13: Ability to identify clock and its function.
Ability to tell time on an hour.

The descriptive data indicate a very stable performance on the number and measurement tasks for three successive sets of incoming first grade students. Also, the post test data suggest that a great deal of growth in performance ability took place during the first grade experience. Given the evidence of stability in performance on entry into first grade, the next step was to apply a factor analysis procedure to the interview data on the 400 students for which follow up information of subsequent mathematics achievement was available.

The factor analysis procedure was used to cluster the 59 performance observations from each interview into a smaller set of independent or orthogonal factors which could subsequently be employed as independent variables in a regression equation for examining the prediction question. The Multiple Analysis Program System (MAPS) used computes a principal factor solution. Factors with eigenvalues of 1.00 or above were retained for a varimax rotation yielding an orthogonal factor structure with the variance of factor loadings maximized. Nine factors were retained. A listing of the factors that the performance variables loaded most heavily on is presented below.

FACTOR 1

Variable	Loading
Read 6	(0.69)
Read 9	(0.72)
Read 7	(0.49)
Read 10	(0.66)
Write 3	(0.16)
Write 8	(0.55)
Write 4	(0.29)
Write 11	(0.51)
Write 9	(0.50)
Write 6	(0.54)
Write 10	(0.55)
Quality Writing	(0.45)
ID Circle	(0.37)

FACTOR 2

Variable	Loading
Heavier	(0.94)
Lighter	(0.95)

FACTOR 3

Variable	Loading
Rote Count from 1	(0.64)
Count back 5	(0.56)
Count back 10	(0.47)
Count back 13	(0.61)
Read 15	(0.68)
Read 18	(0.60)
Write 15	(0.47)
Write 24	(0.57)
ID Triangle	(0.29)
ID Rectangle	(0.36)
Know personal weight	(0.46)

Examination of the clustered variables following the varimax rotation indicated that the small number of variables composing some of the factors make those factors questionable to consider as useful predictors of achievement. Factors two, five, six, seven and possibly eight are in that category. The variables loading on factor one related to performance on reading and writing numbers within an intermediate size number range.

FACTOR 4

Variable	Loading
ID Penny	(0.37)
ID Nickel	(0.73)
ID Dime	(0.67)
ID Quarter	(0.64)
Most Value	(0.38)
Least Value	(0.16)

FACTOR 5

Variable	Loading
Longer	(0.84)
Shorter	(0.80)
ID Clock	(0.42)
Clock Function	(0.30)

FACTOR 6

Variable	Loading
Recognize 3	(0.58)
Recognize 6	(0.75)
Recognize 10	(0.72)

FACTOR 7

Variable	Loading
Tell time 8:00	(0.88)
Tell time 5:00	(0.90)

FACTOR 8

Variable	Loading
Demo 4 blocks	(0.39)
Disc 3 dots	(0.45)
Read 1	(0.37)
Read 2	(0.74)
Read 3	(0.33)

FACTOR 9

Variable	Loading
Count-on from 6	(0.37)
Demo 8 blocks	(0.47)
Demo 10 blocks	(0.46)
Demo 13 blocks	(0.47)
Disc 6 dots	(0.33)
Disc 10 dots	(0.36)
Disc 13 dots	(0.28)
Order 1-7 dots	(0.41)
Order 1-7 numerals	(0.35)
ID square	(0.46)

Numbers of relatively lesser or greater value loaded on the factor eight or three respectively. Factor three variables appeared to relate to performance with numbers of greater value. Counting tc one hundred, counting backward, reading and writing relatively larger numbers, and knowledge of one's personal weight (a relatively large number) loaded on this factor. Factor four loadings were related to performance on the money tasks. Factor eight included a few simple small number performance tasks. The variables loading on factor nine were generally related to rational counting performances. They required students to apply their counting ability in demonstrating a certain number of blocks, discriminating a certain number of dots when faced with a set of choices, and ordering a set of dot cards (1-7) and a set of numeral cards (1-7). In general it would appear that many tasks dealing with various aspects of measurement did not cluster into a factor, or factors, that included a sufficient number of observations to be very meaningful for the prediction question.

The next step taken was to compute multiple correlations and carry out a multiple regression analysis using the nine factors from the interview performance, sex, and two indicators of socio-economic status as dependent or predictor variables, and students' scores on standardized mathematics tests at the first grade level (1976), the second grade (1977), and the third grade (1978) as the independent or criterion variables.

Table 14 presents the results of this procedure. The "R" column indicates the cumulative amount of the relationship between the criterion variable and each successive predictor variable as it is added to the multiple correlation. The R^2 columns indicate the cumulative amount of variance of the criterion variable that can be attributed to the predictor variables as each is successively included in the computation. The "INC" column simple reflects the size of the increment to R^2 of each of the predictor variables. The "P-Value" column indicates the probability that the increment contributed by the specific variable to R^2 is greater than zero, i.e., is not a chance happening.

	1976 MATH				1977 MATH				1978 MATH			
Predictor Variables	R	R^2	INC	P-VAL	R	R^2	INC	P-VAL	R	R^2	INCR	P-VAL
Factor 1	.438	.192	.192	.0001	.446	.199	.199	.0001	.437	.191	.191	.0001
Factor 2	.446	.199	.008	.0302	.446	.199	.000	.8481	.440	.194	.003	.1806
Factor 3	.562	.316	.117	.0001	.625	.390	.191	.0001	.635	.403	.210	.0001
Factor 4	.571	.327	.010	.1415	.646	.417	.027	.0001	.644	.415	.012	.0130
Factor 5	.574	.329	.003	.6366	.650	.422	.005	.1377	.646	.417	.002	.7299
Factor 6	.588	.346	.017	.0017	.657	.431	.009	.0045	.649	.422	.005	.1003
Factor 7	.596	.355	.010	.0116	.659	.434	.002	.1765	.653	.426	.005	.0439
Factor 8	.605	.366	.011	.0798	.663	.439	.005	.2126	.656	.430	.004	.4240
Factor 9	.634	.401	.035	.0008	.681	.464	.025	.0004	.676	.457	.027	.0002
Sex	.637	.406	.005	.0399	.681	.464	.000	.5554	.680	.463	.006	.0051
Current Income Ave Pts.	.641	.412	.006	.0233	.691	.477	.013	.0001	.683	.466	.003	.0410
Rate for Key Indicators	.661	.438	.026	.0001	.697	.485	.008	.0008	.691	.477	.011	.0001

TABLE 14: Multiple correlation (R) and variance (R^2) using 1976, 77, 78 mathematics achievement as criterion.

A few observations can be made by examining Table 14. About 40% (0.401) of the variance on the first grade mathematics achievement test and about 46% (0.464 and 0.457 respectively) of the variance on the second and third grade mathematics achievement tests can be predicted by performance on the entering first grade interviews (cumulative R^2 to Factor 9). When sex and indicators of socioeconomic variables are added, about 44% of the variance on the first grade achievement tests can be predicted and that increases to about 48% on the second and third grade achievement tests.

Factors one, three, and nine make the greatest consistent contributions to the prediction of achievement over the three year period. Combined, they account for about 44% of the variance on the subsequent achievement measures. Factor three, the big number factor, appeared to gain in importance as a predictor of achievement. It contributed about 12% (.117) to the prediction of variance on the first grade achievement test, but by third grade it contributed 21% (.210) to the prediction. Factor four, the money factor, appears to make a small contribution to prediction of the second and third grade achievement scores but not the first grade. Factor six appeared to make a small contribution at the first and second grade but not at third. The sex and socioeconomic variables make only a small contribution to prediction of the achievement score performance. In general it would appear that performance of students on individually administered number tasks on entry to first grade is quite a powerful predictor of subsequent achievement in mathematics as measured by performance on standardized achievement tests. Results of this preliminary analysis indicates that about 45-50% of the variance on subsequent achievement tests can be predicted from interview performance.

The predictive potential of the various number tasks on the interview may offer support for Gelman's (1978) contention that pre-school children's skill at counting is related to their ability to reason numerically. She has presented evidence that children as young as two or three years of age apply "how-to-count" principles. The counting principles constitute a scheme in the Piagetian sense, and as such guide and structure the child's counting behaviors during the pre-school years. As number reasoning through counting develops, the focus of attention can begin to shift from specific numerosities to numerical relationships. About this point in development, a child is able to "conserve" number. It may be that at the point of time most children enter first grade, the counting principles have been in place long enough to assure some stability in performance, yet skill in counting is still developing. Thus students vary considerably in their performance on counting and number tasks allowing for the prediction potential. Readministration of the interviews at the end of first grade indicated that the tasks are by then very easy for the majority of students, thus restricting the variability, and limiting the predictive potential.

Taloumis (1979) has recently reported on the value of Piagetian conservation tasks administered at the first, second, and third grades compared to the value of an intelligence test given at the second grade level or a primary mental abilities test at the kindergarten level in predicting mathematics achievement and teacher's grades in mathematics one, two, three, and four years later. The results seemed to indicate that the conservation tasks were particularly useful for predictive purposes when administered in grade two rather than grades one or three. She observed that the conservation tasks may have been too hard for first graders, and thus fail to predict; in like manner they may be too easy for third graders, and again fail to predict.

It may be that at the stage of dramatic development change taking place about the time youngsters enter school (often referred to as the "six-year change") the predictive power of performance on tasks such as counting may emerge, rise, and then decline over a relatively short period of time.

The initial evidence presented in this study suggests that a set of relatively simple tasks dealing with rote counting, rational counting, and reading and writing numbers, administered to incoming first grade students can be quite a powerful predictor of subsequent achievement in mathematics.

References

Gelman, R. And Gallistel, C. The Child's Understanding of Number. Cambridge, MA: Harvard University Press, 1978.

Lavin, D. The Prediction of Academic Performance. New York: Russell Sage Foundation, 1965.

Taloumis, T. Scores on Piagetian Area Tasks as Predictors of Achievement in Mathematics over a Four-Year Period, Journal for Research in Mathematics Education, 1979, 10, 120-134.

Review of Research: Young Children's Interpretations of Mathematics Pictures

Patricia F. Campbell
Purdue University

Pictures are an important tool used in the teaching of primary mathematics and in diagnostic techniques. Pictures depicting mathematical relationships may serve as a reference to assist students in solving a related problem or they may provide a setting in which a concept may be interpreted. However, the ability to interpret pictures is not innate. As noted by Levie (1973), young children, as compared with educated adults, are "markedly deficient in learning from pictures and the differences appear to be qualitative as well as quantitative". (p. 38) Before interpreting diagnostic measures which utilize pictures or before prescribing instructional techniques involving pictures, it is necessary to consider what influence the pictures may have on the child's mathematical perceptions.

Children's Perceptions of Pictures

When confronted by a picture, what does a child notice? Which features are singled out as key figures and which features are absorbed by the background field? Picture interpretation behavior possesses development characteristics which seem to be related to both maturation and prior learning experience (Higgins, 1978; Randhawa, Back & Myers, 1977).

At around age 3, children can pick out and name recognizable objects in pictures; however, it is easier if the illustration is a frontal, full view rather than a partial, sideways view. Flat, two-dimensional illustrations are easier to interpret as young children do not understand the rules of linear perspective (Tucker, 1978). Initially children perceive discrete items or parts of a picture (frequently insignificant details) to the exclusion of recognition of the whole event being depicted by the picture. Only gradually do the children develop the ability to describe a picture in terms of overt action or to perceive relationships between the illustrated characters (Amen, 1941; Travers, 1969).

Until age 6 or 7 the child prefers simple pictures using strong contour lines without variations in shading or too much overlapping detail (French, 1952; Hutt & McGraw, 1969; Stewig, 1974; Wohlwill, 1975). With more complex pictures involving sketchy, irregular outlines and variations in shading or with linear perspective conventions regarding size and overlaps, 6 or 7 year olds may identify random details but their many small eye movements and use of random visual fixation (Mackworth & Bruner, 1970) make it difficult for them to fuse such details into a comprehensible whole with meaning (Tucker, 1978).

At around the fourth grade children begin to prefer complex illustrations (French, 1952) as increased visual search ability aids them in organizing detail and in drawing inferences from pictures (Tucker, 1978).

Travers and Alvarado (1970) hypothesized that "a child prefers the most complex presentations that he is able to organize perceptually" (p. 59). Note, however, that the familiarity of the depicted items or scenes, as well as the artistic features employed, affect the perceptual complexity of a picture (Higgins, 1978; Travers, 1969). Motion is one of the more difficult representations for children to perceive in static pictures (Brooks, 1977; Campbell, 1976, 1979; Friedman & Stevenson, 1975; Saiet, 1979; Schnall, 1968; Travers, 1969). Also cultural differences affect picture interpretation at all ages (e.g. Deregowski, 1978; Bishop, 1977). There is no one true measure of pictorial perception, rather the question is whether the child's perception is sufficient for the task at hand.

Stages of Picture Interpretation

Observational studies using typical illustrations suitable for young children have noted that children's perceptions tend to develop from enumeration (naming or identifying objects) to description (noting overt activity while still perceiving the objects in isolation) to interpretation (perception of a theme, of relationships between the objects to form a unified whole; making inferences) (Amen, 1941; Carpenter, 1964; Miller, 1938). In repeated timed exposures to the same photographs, young children (nursery and kindergarten children) tended to repeat the same identification over and over while older children (third and sixth graders) tended to expand their perception with each presentation of the photograph (Travers, 1969). Thus the ability to view a picture is not the same as the ability to interpret or "read" a picture.

In recent studies involving first-grade children using mathematics textbook-type illustrations, similar sequential stages have been noted (Campbell, 1976, 1980). The illustrations used in Campbell's studies were either single pictures or sequences of three pictures which portrayed motion by either postural cues (i.e., depicting arms or legs in the state of non-equilibrium) or conventional cues (i.e., depicting clouds of dust or wavy lines about the figures). It is interesting to note that these pictures are usually accompanied by a number sentence in textbooks and the pictures are meant to serve as a reference for solution or as a setting for problem interpretation. However, it is necessary for the children to relate the characters depicted in the pictures and to see the action portrayed before they can associate the pictures with the addition or subtraction of whole numbers. Yet these are the same picture interpretation skills which psychological studies indicate that first-grade children may be developing, but have not yet obtained.

The stages of mathematics picture interpretation noted by Campbell were observed in individual interviews with 192 first-grade children in five elementary schools (three in Lafayette, Indiana; two in Tallahassee, Florida) during late February and early March in two consecutive years. More recently 60 more first-grade children were interviewed from a sixth school in Indiana. The stages of interpretations were again substantiated. Although the children were using three different major elementary mathematics textbook series, no significant relationship between the children's interpretations and the adopted series has been noted.

The technique used in all the interviews was to present each child with a picture or sequence of three pictures and to tell the child that the picture(s) told a story. The children were then asked either to tell the story told by the picture(s) or to use toys (corresponding to the characters in the picture(s)) to show the story. This procedure was repeated for ten illustrations (five addition, five subtraction). All children who were asked to show the story with the toys at the same time told a story to explain their placement of the toys without any prompting from the interviewer. These stories did not differ from those told by children who did not have the toys present.

The stories told by the children were in one of four stages or categories. At the first stage the stories were general in nature, elaborating on one character or item in the picture(s) without concern as to the theme of the illustration or the mathematical concept being portrayed. The length and complexity of these stories seemed bounded only by the children's imagination and their knowledge of facts about the type of character which they recognized in the illustration. At the next level, the stories revealed perception of disjoint sets, but neither perception of a mathematical relationship between the sets nor perception of motion was indicated. For these children, the desire to count or group the characters was dominant with the emphasis being on enumerating the members of the disjointed sets. Stories in the third category described disjoint sets and indicated recognition of motion, but did not relate the sets. The listing of characters was still dominant with motion serving as an additional feature to be noted in the description of the separate sets, rather than as an indicator of a relationship between the sets. The fourth level of stories described a total set formed by the relationship of the two disjointed sets with an action either joining or separating the sets.

Although many children consistently told similar stories for all ten items reflecting either stage 1, 2, 3 or 4, there were a number of children who seemed to be in a state of transition between stages 2 and 3 or between stages 3 and 4.

Pictures are used not only as a communication medium in primary textbooks, but they are also utilized as a means of aiding young children in making the transfer from concrete, manipulative settings to abstract, symbolic representations. However, it is first necessary for the child to interpret the pictures. The stages of picture interpretation noted here were verified for pictures which depicted a theme or story. However, Travers (1969) and Ames et al. (1953) have noted similar levels with other forms of illustrations. Before using a picture to instruct or evaluate, it is necessary to determine what the picture conveys to the child. Recently Silliman (1979) coined the term "representational competence" as the "capacity to understand the equivalence between verbal and graphic/ikonic modes of representation irrespective of whether the media is verbal or graphic/ikonic" (p. 387). She noted that when language comprehension tasks were embedded within a picture it was possible that the result was an increase in conceptual complexity, particularly for the language impaired or learning-disabled child. It is quite plausible that a similar increase in conceptual complexity occurs when mathematical tasks are presented within pictures.

Characteristics of Pictures and their Effect on Interpretation

Little is known concerning which characteristics of pictures influence interpretation and hence which forms of pictorial representation best convey the intended concepts to the learner. Much of the research in psychology concerning learning from pictures is focused upon comparisons with learning from words. Many of these laboratory studies are based upon paired-associate tasks or recall (recognition) tasks. Different hypotheses have subsequently been offered concerning the processing of verbal and visual information within either a single or a dual system (e.g. Paivio, 1971), with substantial evidence favoring the dual model of two separate memory systems, one to process verbal symbols and one to process nonverbal visual images.

Pictorial materials vary along many dimensions. Fidelity is the degree to which a picture provides a view with "sensory imput" similar to that which he would have if he were viewing the scene depicted by the picture (Travers & Alvarado, 1970). Another dimension is the presence or absence of background settings in which the central items are presented. The amount of detail (color, line drawing, shading and photograph) and the use of three-dimensional cues are other characteristics which may vary between pictures.

During the 1940's and 1950's a group of theories, referred to by Dwyer (1967) as the "realism theories", were based upon the assumption that the more realistic the instructinal device, the more effective it will be as an aid for learning. These theories held that more realistic materials present more visual cues to the learner and therefore provide more information for understanding. In contrast a separate group of researchers labelled the "relevant cue" theorists (Parkhurst, 1975) contended that the learner may be flooded with too much visual information and that those elements of an illustration which convey little pertinent information may be edited without impairing instructional value. There are supporting studies (Dwyer, 1972; Travers, 1969) for the "relevant cue" theories.

Although children and adults generally prefer color illustrations (Bloomer, 1960; Hildreth, 1936; Rudisill, 1952) over black-and-white pictures, the effectiveness of color in instructional settings is not clear. However, recent evidence (Dwyer, 1976) seems to indicate that for older students color may cause a sufficient increase in interest which results in a significant increase in learning, as compared to similar black-and-white illustrations.

Many illustrations in elementary school textbooks are stylized representations. These are pictures which use adopted codes to portray features which, as judged by the code, represent the real world although they deviate from the true optical code. Cartoon-like drawings of personified animals or objects are examples of stylistic representations. It has been demonstrated that, in mathematics textbook-type illustrations, first grade childern interpret stylistic representations with an ease equivalent to their interpretations of realistic line drawings of children. Hence, use of both types of illustrations may provide variety without adversely affecting understanding (Campbell, 1978).

Another stylized code typically presented in textbook illustrations is the use of motion cues. Most textbook illustrations require the student to convert the still picture into a dynamic scene in order to derive all the information intended by the illustrator. A static picture must be viewed as an instance which has emerged from previous events and as an instance which is leading up to or causing new events. Children who are still viewing pictures in a descriptive or enumerative mode are probably unable to reason antecedent or subsequent events (Higgins, 1978); many children are 9 or 10 years of age before they provide such an interpretation with inferences concerning meaning or significance (Tucker, 1978). It has been shown that a sequence of three pictures to depict an event may prompt interpretation at a higher level than a single picture of the event. In addition, viewing and interpreting sequences causes learning which aids later interpretation of single pictures (Campbell, 1979).

Three conventions are typically employed to depict motion in static pictures; to depict the moving object as a blurred form; to represent the moving object in a state of non-equilibrium; or to draw vibration or wake lines behind the moving object. In their study of preschool, first-grade, sixth-grade and college students' interpretatons of motion cues, Friedman and Stevenson (1975) noted that the two younger groups of children did not classify pictures with the blurred, wake or vibration lines as "moving" as often as they did the pictures involving the postural cues. Similarly Brooks (1977) asserted that second-grade children could not use vibration and wake lines as a facilitator in a recall task using cartoon pictures. In contrast, Campbell (1976) noted that first-grade children were more likely to observe motion in pictures depicting both postural and conventional cues than in pictures depicting only postural cues. Because the conventional cue figures in Friedman and Stevenson's study had straight limbs visible, there may have been a confounding effect if the children interpreted the straight limbs as a postural cue indicating no movement. A subsequent comparison of postural cues (limbs in non-equilibrium) and conventinal cues (wake or vibration lines about figures without limbs such as fish or snails) by Campbell (1979) indicated that motion cues did have a significant effect on young children's interpretations. Postural cues were more readily interpreted, but first-graders were able to "see motion" via conventional cues. A related study by Saiet (1979) with 4, 7 and 11 year olds supported the studies by Campbell. Saiet's data indicated that given a human figure with postural cues, all age groups identified motion with or without the addition of conventional cues, although the addition of the conventional cues increased the level of motion perception. Given a human figure with straight limbs, the 4 year olds did not perceive motion with the addition of conventional cues. However, about 50% of the 7 year olds and 75% of the 11 year olds perceived motion when vibration or wake lines were added to the straight limb figure. Blurring was typically perceived as a shadow or form behind the figure, rather than as a motion cue, by all age groups in Saiet's study.

The Use of Pictures in Instructional Settings

Research reports of the relationship between pictures and mathematics learning have considered the effect of pictorial representations on children's ability to solve word problems. Sherrill (1971) noted that written problems with accompanying accurate illustrations were more readily solved than either the problems only or the problems with inaccurate illustrations. However, significant correlations between problem solving, IQ, reading level and mathematics grades were noted which may affect generalization. In a similar study involving diagrams, Nelson (1975) noted that diagrams could be used as an effective technique in problem solving. Kulm et al. (1974) noted that while no difference was noted within average or high IQ groups, the low (below 110) IQ group was able to solve significantly more textbook problems without accompanying pictures than problems with pictures. Botsmanova (1960) noted that illustrations which portrayed the mathematical structure of a problem aided solution whereas illustrative pictures which portrayed only the objects noted in the problem, but not their mathematical relationships, were not useful for problem solving. Hence, these studies seem to indicate that pictures may be a useful tool in problem solving; however, the extent and manner of their use is not clear.

There are a number of studies which have considered the effect of pictures as an adjunct to the understanding of written or oral text within the reading curriculum. These studies may supply insight as to the effect of pictures accompanying elementary mathematics lessons as in both situations the illustrations are intended to aid the student in interpreting accompanying symbolic material. As such these studies may suggest hypotheses for research concerning the role of pictures in learning elementary school mathematics.

The influence of pictures with reading has been investigated within two settings, learning to read or decode written words and learning to read passages for comprehension. Another area of study is the influence of visual imagery on recall and comprehension of written prose (e.g. Levin & Divine-Hawkins, 1974). This area will not be reviewed in this paper.

Samuels (1967, 1970) noted that under experimental conditions kindergarten children acquired sight vocabulary of four words faster under a word-only condition than under a simultaneous picture-word condition. In a subsequent study with first-grade children he noted that while pictures did not influence better readers, the presence of an illustration seemed to interfere with the acquisition of sight vocabulary among poorer readers. He concluded that pictures may be acting as distractors and thereby retard acquisition of reading skills. This focal attention hypothesis has subsequently become the subject of many investigations with varying results (e.g. Arlin, Scott & Webster, 1978-79; Harzem, Lee & Miles, 1976; Montare, Elman & Cohen, 1977; Samuels, 1977; Willows, 1978). The impact of pictures and decoding written text is not clear.

Conversely research concerning pictures and comprehension of oral prose seems to indicate that pictures do facilitate prose learning. These studies orally presented fictional narratives (less than 12 sentences) to young children (usually via an audio-tape) with either the presence or absence of pictures overlapping the story content. All learning was demonstrated by factual recall in posttests and/or retention tests (e.g. Guttman, Levin, Pressley, 1977; Lesgold, DeGood, Levin, 1977; Lesgold, Levin, Shimron & Guttman, 1975; Levin, Bender & Lesgold, 1976; Levin & Lesgold, 1977). Oral prose was utilized in order to avoid the confounding of decoding (sight reading) with comprehension. It is interesting to note that while the focal attention hypothesis would contend that accompanying pictures would divert the readers' attention from the written text, the distraction need not generalize to an orally presented text as listening and looking in this setting may be complementary rather than competing information-processing activities (e.g. Levin & Divine-Hawkins, 1974).

Individual Differences

The influence of individual difference variables on picture interpretation is not clear. Recently the need for such investigation has been noted (Ausburn & Ausburn, 1978; Parkhurst, 1975). As a result a study was undertaken in early April, 1979, to investigate the possible relationship between mathematics picture interpretation skills, spatial visualization and field-dependence-independence. It had previously been noted that neither mathematics achievement (as measured by a standardized diagnostic test) nor cognitive development (as measured by Piagetian conservation tasks) seemed to be related to mathematics picture interpretation skills (Campbell, 1978, 1979).

It was hypothesized that the ability to perceive motion and to see relationships between the characters in dynamic illustrations may be related to spatial ability. When interpreting a dynamic mathematics textbook picture, the children must use mental images to reason what event occurred prior to the depicted scene and what event immediately followed the scene. They must also interpret static two-dimensional features within the code by an artistic convention which deviates from the real optical code. Because spatial visualization refers to the formation of visual images of objects and subsequent movement or change in the objects or their properties, this factor of spatial ability was used in the study as an independent variable.

It was further hypothesized that picture interpretation may be related to the student's field-dependence or field-independence cognitive style (Kreider, 1977). In particular, when presented with a mathematics picture, the student must restructure the picture from a mathematical perspective, rather than simply adhering to the organization presented artistically. Field-dependent and field-independent individuals differ markedly in their restructuring ability, in both perceptual activities and intellectual functioning. Field-independent people seem better able to restructure their initial perception to obtain the interpretation demanded by the constraints of the problem. Field-dependent people adhere to the perceptually compelling interpretation (Case & Globerson, 1974; Nebelkopf & Dreyer, 1970; Witkin, Moore, Goodenough & Cox, 1977; Witkin & Goodenough, 1977).

Preliminary analysis indicates that there is a significant relationship between picture interpretation skills and field-dependence-independence as revealed both in the students' stories and in their characterization of the picture by number sentences. An increase in field-dependence-independence scores is correlated with an increase in picture interpretation scores. No significant relationship between spatial ability and picture interpretation was noted. No significant sex differences were noted. However, there are significant differences between the interpretation of addition and subtraction pictures.

Conclusions

A review such as this seems to raise more questions concerning children's interpretations of mathematics pictures than it answers. However, some considerations for instruction and evaluation may be offered. Childern interpret pictures differently from adults. Generally, they obtain less information from pictures, both in the quantity and quality of their observations. Young children frequently overlook, ignore or "incorrectly" perceive the intent of a picture. Before interpreting diagnostic measures which utilize pictures or before prescribing instructional techniques involving pictures, it is necessary to consider what influence the pictures may have or have had on the child's reported mathematical perceptions. Ask the child to tell what story the picture is conveying or to use objects to enact the illustration. Discuss the relationship between the mathematical concept and the contents of the picture. Do not assume a transfer from concrete objects to pictures or from pictures to numerals. The transfer must be taught. Use sequences depicting antecedent and subsequent events to aid interpretation of single pictures. Illustrations may aid students in understanding mathematical concepts, but only if they first understand the pictures.

References

Amen, E.W. Individual differences in apperceptive reaction: A study of the response of preschool children to pictures. Genetic Psychology Monographs, 1941, 23, 319-385.

Ames, L.B., Learned, J., Metraux, T. & Wheeler, R. Development of perception in the young child as observed in responses to Rorschach test blots. Journal of Genetic Psychology, 1953, 82, 1983-204.

Arlin, M. Scott, M. & Webster, J. The effects of pictures on rate of learning sight words: A critique of the focal attention hypothesis. Reading Research Quarterly, 1978-79, 14, 654-660.

Ausburn, L.J. & Ausburn, F.B. Cognitive styles: Some information and implications for instructional design. Educational Communication and Technology: A Journal of Theory, Research and Development, 1978, 26,, 337-354.

Bishop, A. Is a picture worth a thousand words? Mathematics Teaching, 1977, No. 81, 32-35.

Bloomer, R.H. Children's preference and responses as related to styles and themes of illustrations. The Elementary School Journal, 1960, 60, 334,340.

Botsmanova, M.E. (The forms of pictorial visual aids in instruction in arithmetic problem solving.) In J.W. Teller (Trans.) & J. Kilpatrick & I Wirszup (Eds.), Soviet Studies in the Psychology of Learning and Teaching Mathematics (Volume 6). Pasadena, California: A.C. Vroman, 1972. (Reprinted from Reports (Doklady) of the Academy of Pedagogical Sciences of the RSFSR, 1960, 5.)

Brooks, P.H. The role of action lines in children's memory for pictures. Journal of Experimental Child Psychology, 1977, 23, 98-107.

Campbell, P.F. The Role of Pictures in First Grade Children's

Perception of Mathematical Relationships (Tech. Report No. 8). Tallahassee, Florida: The Florida State University, Project for the Mathematical Development of Children, 1976. (ERIC Document Reproduction Service No. ED 144 807)

Campbell, P.F. Textbook-type pictures and first-grade children's perception of mathematical relationships. Journal for Research in Mathematics Education, 1978, 9, 368-374.

Campbell, P.F. Artistic motion cues, number of pictures and first-grade children's interpretatin of mathematics textbook pictures. Journal for Research in Mathematics Education, 1979, 10, 148-153

Campbell, P.F. What do children see in mathematics textbook pictures? Arithmetic Teacher, 1981, 28, 5, 12-18.

Carpenter, H.M. Study skills: Developing picture-reading skills. The Instructor, 1964, 74, 37-38; 130.

Case, R. & Globerson, T. Field-independence and central computing space. Child Development, 1974, 45, 772-778.

Deregowski, J.B. Pictures as a means of communication: A psychologist's view. Educational Broadcasting International, 1978, 11, 7-12.

Dwyer, F.M. Adapting visual illustrations for effective learning. Harvard Educational Review, 1967, 37, 250-263.

Dwyer, F.M. The effect of IQ level on the instructional effectiveness of black-and-white and color illustrations. AV Communication Review, 1976, 24, 49-62.

Dwyer, F.M. A Guide for Improving Visualized Instruction. University of Pennsylvania: Learning Services, 1972.

Friedman, S.L. & Stevenson, M.B. Developmental changes in the understanding of implied motion in two-dimensional pictures. Child Development, 1975, 46, 773-778.

French, J.E. Children's preferences for pictures of varied complexity of pictorial pattern. The Elementary School Journal, 1952, 53, 90-95.

Guttman, J., Levin, J.R. & Pressley, M. Pictures, partial pictures and young children's oral prose learning. Journal of Educational Psychology, 1977, 69, 473-480.

Harzem, P., Lee, I. & Miles T. The effects of pictures on learning to read. British Journal of Educational Psychology, 1976, 46, 318-322.

Higgins, L.C. A factors analytic study of children's picture intrepretation behavior. Educational Communication and

Technology: A Journal of Theory, Research and Development, 1978, 26, 215-232.

Hildreth, G.H. Color and picture choices of young children. Journal of Genetic Psychology, 1936, 49, 427-435.

Hutt, C. & McGraw, P. Do children really favor visual complexity? Psychonomic Science, 1969, 17, 113-114.

Kreider, M.S. An investigation of the relationship of perceptual field-dependence-independence, the mode of representation used in instruction, and the ability to recognize signs of nutritional deficiency (Doctoral dissertation, University of Maryland, 1976). Dissertation Abstracts International, 1977, 37, 6909A. (University Microfilms No. 77-10, 280)

Kulm, G., Lewis, J.F., Omari, I. & Cook, H. The effectiveness of textbook, student-generated and pictorial versions of presenting mathematical problems in ninth-grade algebra. Journal for Research in Mathematics Education, 1974, 5, 28-35.

Lesgold, A.M. , DeGood, H. & Levin, J.R. Pictures and young children's prose learning: A supplementary report. Journal of Reading Behavior 1977, 9, 353-360.

Lesgold, A.M., Levin, J.R., Shimron, J. & Guttman, J. Pictures and young children's learning from oral prose. Journal of Educational Psychology, 1975, 67, 636-642.

Levie, W.H. Pictorial research: An overview. Viewpoints, 1973, 49, 37-45.

Levin, J.R., Bender, B.G. & Lesgold, R.M. Pictures, repetition and young children's oral prose learning. AV Communication Review, 1976, 24, 367-380.

Levin, J.R. & Divine-Hawkins, P. Visual imagery as a prose learning process. Journal of Reading Behavior, 1974, 6, 23-30.

Levin, J.R. & Lesgold, A.M. On Pictures in Prose (Theoretical Paper No. 69). Madison, Wisconsin: The University of Wisconsin, Wisconsin Research and Development Centre for Cognitive Learning, December, 1977. (ERIC Document Reproduction Service No. ED 153 184)

Mackworth, N.H. & Bruner, J.S. How adults and children search and recognize pictures. Human Development, 1970, 13, 149-177.

Miller, W.A. What children see in pictures. The Elementary School Journal, 1938, 39, 280-288.

Montare, A., Elman, E. & Cohen, J. Words and pictures: A test of Samuels' findings. Journal of Reading Behavior, 1977, 9, 269-285.

Nebelkopf, E.B. & Dreyer, A.S. Perceptual structuring: Cognitive style differences in the perception of ambiguous stimuli. Perceptual and Motor Skills, 1970, 30, 635-639.

Nelson, G.T. The effects of diagram drawing and translation on pupils' mathematics problem-solving performance. (Doctoral Dissertation,University of Iowa, 1974) Dissertation Abstracts International, 1975, 35, 4149A. (University Microfilms No. 75-1234).

Paivio, A. Imagery and Verbal Processes. New York: Holt, Rinehart & Winston, 1971.

Parkhurst, P.E. Generating meaningful hypotheses with aptitude-treatment interactions. AV Communication Review, 1975, 23, 171-183.

Randhawa, B.S., Back, K.T. & Myers, P.J. Visual Learning. Paper presented at the annual meeting of the Association for Educational Communications and Technology, Miami Beach, Florida, April 26, 1977. (ERIC Document Reproduction Service No. ED 143 319)

Rudisill, M. Children's preferences for color versus other qualities in illustrations. Elementary School Journal, 1952, 52, 444-457.

Saiet, R.A. Children's Understanding of Implied Motion Cues. Paper presented at the annual meeting of the Association for Educational Communications and Technology, New Orleans, Louisiana, March 1979. (ERIC Document Reproduction Service No. ED 172 801)

Samuels, S.J. Attentional processes in reading: The effect of pictures on the acquisition of reading responses. Journal of Educational Psychology, 1967, 58, 337-342.

Samuels, S.J. Can pictures distract students from the printed word: A rebuttal. Journal of Reading Behavior, 1977, 9, 361-364.

Samuels, S.J. Effects of pictures on learning to read, comprehension and attitudes. Review of Educational Research, 1970, 40, 397-407.

Schnall, M. Age differences in the integration of progressively changing visual patterns. Human Development, 1968, 11, 287-295.

Sherrill, J.M. The effects of differing presentations of mathematical word problems upon the achievement of tenth grade students (Doctoral dissertation, The University of Texas at Austin, 1970). Dissertation Abstracts

International, 1971, 31, 3427A. (University Microfilms No. 71-191).

Silliman, E.R. Relationship between pictorial interpretation and comprehension of three spatial relations in school-age children. Journal of Speech and Hearing Research, 1979, 22, 366-388.

Stewig, J.W. Children's picture preference. Elementary English , 1974, 51, 1012-1013.

Travers, R.M.W. A Study of the Advantages and Disadvantages of Using Simplified Visual Presentations in Instructional Materials. Final Report. USPE Grant No. OEG-1-7070144-5235, 1969 (ERIC Document Reproduction Service No. ED 031 951).

Travers, R.M. W. & Alvarado, V. The design of pictures for teaching children in elementary school. AV Communication Review, 1970, 18, 47-64.

Tucker, N. Ways children see. Times Educational Supplement (London). No. 3305, November 3, 1978, 21.

Willows, D.M. Individual differences in distraction by pictures in a reading situation. Journal of Educational Psychology, 1978, 70, 837-947.

Witkin, H.A. & Goodenough, D.R. Field Dependence Revisited (Research Bulletin No. 77-16). Princeton, New Jersey: Educational Testing Service, December, 1977.

Witkin, H.A., Moore, C.A., Goodenough, D.R. & Cox, P.W. Field-dependent and field-independent cognitive styles and their educational implications. Review of Educational Research, 1977, 47, 1-64.

Wohlwill, J.F. Children's voluntary exploration and preferences for tactually presented nonsense shapes differing in complexity. Journal of Experimental Child Psychology, 1975, 20, 159-167.

Locus of Control and Feedback in Diagnostic Mathematics Testing Under Conditions Promoting Initial Success or Failure

Jon M. Englehardt
Department of Elementary Education
Arizona State University

R. Keith Van Wagenen
Department of Educational Psychology
Arizona State University

Noel Thomas
Mitchell College of Advanced Education
Australia

This was an investigation of a possible interaction between the learner characteristic of locus of control and knowledge of response (correct - incorrect) feedback on a diagnostic mathematics test. Children were asked to solve two series of test items, some children receiving feedback and some not. One series represented initial success progressing toward failure, the other initial failure progressing toward success.

In testing situations, telling students whether their answers are correct or not is commonly prohibited or delayed until students have completed the test (e.g., Anastasi, 1976; Cronbach, 1960). Finding little evidence concerning knowledge of response (correct - incorrect) feedback in diagnostic-type testing situations, Engelhardt (1978) reported an investigation of whether item-by-item knowledge of response feedback influenced students' performance on a paper-and-pencil mathematics test. The results suggested that fourth and sixth-grade students receiving such feedback did not perform differently than students receiving no feedback. In earlier investigations, Betz (1977) found such feedback to lead to higher verbal ability performance for both high and low ability subjects, Rust, Strang and Bridgeman (1977) found no difference in seventh-graders' science test performance to result from such feedback, and Gilmer (Note 1) found college students receiving feedback to perform significantly lower than those receiving no feedback. Clearly, the issue remains open.

In none of these studies were item arrangements or intrapersonal factors like anxiety or locus of control taken into account. Yet both item sequencing and individual difference variables have been identified as potentially influencing diagnostic mathematics testing (Engelhardt, 1976, 1978). Considering factors such as these may better explain the impact of knowledge of response feedback on diagnostic testing.

When items are ordered in ascending (easy-to-hard, simple-to-complex) fashion, the individual presumably experiences initial success (i.e., avoids initial failure), thereby minimizing anxiety and maximizing performance. Unfortunately, since the individual may respond to many items which might be considered "easy", it may be argued that the individual could, (a) lose concentration or be bored and fail to perform as the items become more difficult or (b), noting the progressively increasng difficulty of the items, "psych out" before getting to the appropriate performance limit -- thus giving artificially low performance. When items are ordered descendingly from hard-to-easy, it is usually argued that the individual wastes little time with items on which he/she cannot perform, proceedng quickly to items at an appropriate level of difficulty. The counterargument is that through initial failure on items the individual generates excessive anxiety and, in moving quickly past the "too difficult" items, may only begin seriously attempting items below that person's ability -- again giving an artificially low performance.

In previous research when items have been arranged in ascending order of difficulty, performance has been higher than when other formats were presented (e.g., Brenner, 1964; Sax and Carr, 1962; Sax and Cromack, 1966), and descending arrangements have led to lower performance (Hambleton, Note 2; Towle and Merrill, Note 3). However, other findings continue to appear. Klosner and Gellman (1973) found no difference in the performance of college students for different item arrangements, except that low achieving students did least well with an ascending arrangement. Suggesting performance differences might more be the result of limited administration time, Marso (1970) found when no time limit was imposed that college students' test performance was unaffected by item arrangement. Shoemaker (1970) reported no significant effect for item order, but when students were tested one item at a time, they performed best on the ascending arrangement. Zeineh, Pierce and Silverman (Note 4), in an investigation with fifth-grade students, found almost identical performance on ascending and descending arrangements of computation items from the Metropolitan Achievement Test. Although best performance was on a random arrangement, there was a slight trend for high test-anxious students to perform better on the ascending arrangement. None of the studies reviewed showed a consistent relation between item arrangement, anxiety and test performance. Item arrangement is clearly not a black-and-white issue.

Locus of control, the attribution of credit/blame for success and failure to self or others, may be states of perception which help explain performance after feedback which is heavily positive or heavily negative. Weiner (Weiner, Heckhausen, Meyer and Cook, 1972) proposed two dimensions for helping explain achievement-related behavior -- internality/externality (attribution of success/failure to self or others/outside forces) and stability of the attribution over time. Ability and task difficulty are considered stable attributions, since they would not vary if the task were reattempted, but luck and effort are unstable, since they might change on a second attempt. Bar-Tal (1978) reviewed the literature on attribution theory as it relates to achievement.

Individuals who attribute their success to internal causes (ability or effort) are likely to be encouraged by positive feedback and, compared to those not getting feedback, perform better when items are arranged so as to maximize initial success, i.e., ascending item order. Those attributing their failure to stable internal causes (lack of ability) would probably become discouraged by negative feedback showing failure, i.e., descending item order. However, when failure is attributed to unstable internal causes (lack of effort), such individuals might try harder, become less discouraged and persist despite negative feedback, perhaps performing better than those receiving no feedback, when items are arranged to emphasize initial failure, i.e., descending item order.

Those attributing success and failure to external causes (task difficulty or luck) likely see their performance as unrelated to their effort and beyond their control. As a result feedback would probably have little impact on subsequent performance, and such individuals should perform similarly to those not receiving feedback, regardless of the item arrangement.

Subjects

The sample consisted of 36 third-grade students, representing two classrooms in a parochial school in suburban Phoenix, Arizona. Subjects were randomly selected from each classroom in a stratified sample so as to represent different mathematics achievement levels. High achievers were those who performed more than 3 months above grade level, low achievers more than 3 months below grade level, and middle achievers between these extremes.

Materials

Two versions,ascending and descending, of the addition and subtraction subscales of the Key Math Diagnostic Arithmetic Test (Connolly, Nachtman and Pritchett, 1971) were used to measure subjects' mathematics performance. These two subscales were selected for two reasons. First, all students should have been able to perform to some degree on the items, regardless of their mathematics achievement level. Second, as per the grade equivalents, children in the third grade should on the average perform correctly on about half of the subscale items; thus, whether approached in an ascending or descendng fashion, the students should encounter about half of the items. The addition subscale consisted of 15 items and subtraction, 14 items.

In both versions of the subscales, students were initially given the first six items from the numeration subscale; these were used as a warm-up and to introduce the flip-card format for some of the later items. Subjects received most items individually and responded to each item on a (5 x 8 inch) page. For subjects who would receive feedback the words "correct" and "incorrect" were printed on their response sheets.

Subjects saw/heard one item at a time, proceeding to the next item only when they had completed an item or indicated they were unable to complete it. This arrangement precluded the child working items out of order.

In the ascending version, items for the subscales were separately administered and arranged in the usual easy-to-hard order. When three items in a row were answered incorrectly, no further items on that sub-scale were administered. The subscale was scored by counting the total number of items answered correctly. In the descending version, items were administered in the reverse, hard-to-easy order. When three items in a row were answered correctly, no further items were administered, and the subscale was scored by counting the total number of items answered correctly plus the easier items not administered. Discounting the possible effects of intervening factors like anxiety, these two versions should result in roughly the same score for an individual.

The measure of locus of control was specially constructed by one of the investigators. Lefcourt (1976) stated:

> If one wishes to use the perception of control as a powerful predictor, then it will most always be profitable to design one's own assessment devices for the criterion of interest. This is similar to stating that people are not so much to be characterized as internals and externals as they may be said to hold internal and external control expectancies about different aspects of their lives... To summarize this point, one would always do well to add an "of what" after the phrase perception of control. (p.153-154)

This point was empirically supported by Butkowsky (Note 5). The measure was specifically constructed to get an indication of subjects' perception of feedback in classroom achievement situations as resulting from internal or external causes and its stability. For example, when a teacher indicates to a child that his/her response is incorrect, this may trigger an attribution for that feedback, such as lack of effort, bad luck, lack of ability, the difficulty of the question, or perhaps even a contrary teacher. Crandall, Katkovsky and Crandall (1955) have successfully designed a locus measure for intellectual achievement; however, it is not restricted to classroom-type achievement and does not distinguish the stability of the attribution.

The items in the locus measure are achievement-feedback specific, paper-and-pencil items, involving a structured response to hypothetical/recalled situations. The items are grouped into three parts. The first six items pertain to classroom achievement in general, the next four to non-classroom achievement, and the last six to classroom mathematics achievement. The four non-classroom achievement items were included to form a break between the other two parts and to help accentuate that one group pertained to mathematics and the other did not. Half of each part of the test involved positive feedback situations and the other half negative feedback situations. Furthermore, four items in each six-item part dealt with positive/negative feedback in convergent situations (i.e., those involving a single, expected response) and the other two with divergent situations (i.e., those involving novel responses and approaches).

Each item included four response choices, from which the subject was to select the one which best reflected his/her beliefs in that situation. Following Weiner's dimensions of causation (Weiner, Heckhausen, Meyer and Cook, 1972), two of the response choices were designed to monitor internal attributions, and two response choices external attributions. One of the choices in each pair suggested a stable attribution and the other an unstable attribution. Figure 1 presents, in general, the four attributions represented by the four response choices.

See Figure 1 on page 63

Scoring on the measure was derived from only the two classroom achievement sets of items. Figure 2 presents the 54 possible scores which can be derived from this measure. The maximum score was doubled when the two achievement situations were combined. Scores marked with an asterisk are those of concern in the present investigation.

See Figure 2 on page 63

Procedure

Prior to any mathematics testing, all students in the two classrooms were administered the locus of control measure. The testing was conducted by an investigator in each classroom separately, without the regular classroom teacher, and students were told that the teacher would not see their responses. Students had written copies of the items, but they were also read aloud by the tester. Both classrooms were tested the same day.

During the next week students from each classroom were randomly selected and assigned to one of four feedback/item-order treatments. These were assigned so as to have about equal representation for the three mathematics achievement levels. Two of the four treatment groups were to receive item-by-item knowledge of response feedback, while the other two received none. All treatment groups, after the six warm-up items, responded to subscale items twice -- once arranged in ascending fashion and once in descending fashion. Both subscales were administered to each subject in one order before they were given in the other order. Difficulty order was counterbalanced in the treatment groups.

In the two treatment groups receiving feedback, subjects proceeded from one item to the next only after the tester had marked the appropriate word, "correct" or "incorrect" on the response sheet. In the ascending item arrangement, when three items in a row were incorrect, the subject was instructed to close the test booklet. This was done when three items in a row were correct for the descending item arrangement. The two groups receiving no feedback were handled exactly the same, but no feedback response was made to the child. Except for the few flip-card items, the three subjects in each group were arranged in a circle, facing outward; in this way they had little idea about the feedback others received or who had finished first. This seating arrangement was intended to roughly equate one-to-one diagnostic testing.

All testing was done by three testers in an empty classroom. Each tester worked with four groups, one of each treatment. Testing was completed in two successive afternoons.

Results

The means and standard deviations for performance on the locus measure are presented in Table 1. In general, subjects tended to be more internal than external in their locus for feedback in classroom-related achievement situations -- accepting more of both praise and blame for the feedback they received. They were more internal for success than failure, with both being dominated by unstable attributions (i.e., effort). Also reported in Table 1 are tests comparing responses on the general achievement and mathematics achievement parts of the measure. Although some differences are apparent between the two parts, overall response was quite similar. Analysis of individuals, case by case, reconfirmed this observation.

In Table 2 are the results of an analysis of variance comparing locus perceptions of subjects receiving feedback with those receiving no feedback. Clearly, there was no difference in perception on any of the specific locus subscales.

See Tables 1 and 2 on page 64

Table 3 presents the means and standard deviations for diagnostic mathematics test performance. Three observations can be made. First, those who received items in an ascending arragement, followed by a descending arrangement (Group 1) tended to perform better on both arrangements than those receiving the item arrangements in reverse order (Group 2). Second, subjects tended to perform better on the descending item arrangement, regardless of which item arrangement they received first. Although these trends were not always statistically convincing, they were quite consistent, whether subjects received feedback or not. Third, subjects did not tend to perform any better on the first version they received than on the second.

See Table 3 on page 65

Disregardng which version of the mathematics test was received first, subjects receiving feedback performed significantly higher on the subtraction subscale than those receiving no feedback. (See Table 4.) This was true whether the subscale items were administered in ascending or descending order. Little difference in performance was apparent on the addition subscale.

See Table 4 on page 65

The thesis of this investigation was that locus of control influences the effect item-by-item knowledge of response feedback has upon mathematics test performance and that this effect becomes particularly apparent when test items are arranged so as to maximize initial success or failure. Specifically the following hypotheses were posed:

(1) Individuals tending to attribute their success to internal causes (ability or effort) are likely to be encouraged by feedback and therefore, compared to those not receiving feedback, perform better when items are arranged so as to maximize initial success, i.e., ascending item order.

(2) Individuals tending to attribute their failure to stable internal causes (lack of ability) are likely to be discouraged by negative feedback and perform more poorly when items are arranged so as to maximize initial failure, i.e., descending item order.

(3) Individuals tending to attribute their failure to unstable internal causes (lack of effort) are likely to try harder and become less discouraged by negative feedback and thus perform better when items are arranged so as to maximize initial failure, i.e., descending item order.

(4) Individuals tendng to attribute their success or failure to external causes (task difficulty or luck) are likely to see their performance as unrelated to their effort and beyond their control, and thus feedback should have little impact on their performance, regardless of item arrangement.

To test these hypotheses, the Johnson-Neyman technique (Ceurvorst, Note 6) was used. Like analysis of covariance, this technique helps determine if, when the effects of a control variable are taken into account, significant group differences exist. However, the Johnson-Neyman technique does not assume homogeneity of regression between groups and, unlike covariance, indicates the region in the space of the covariate where those differences occur.

Information on the regression of the pertinent locus of control variables on mathematics test performance is shown in Table 5. The *F*-ratios reflect the results of tests for homogeneity of regression.

See Tables 5 and 6 on pages 66 and 67

Examination of the scatterplots of residuals for each regression line suggested no obvious curvilinearity. Johnson-Neyman procedures were therefore conducted. The results of the pertinent comparisons of regression lines for regions of significant difference (.05 level) appear in Table 6. Because conclusions relative to scores beyond those actually observed are tenuous, the actual range of scores has been included. These data suggest that, compared to those not receiving feedback, (a) those receiving feedback and who scored roughly between 4.0 and 5.5 (out of 6) on the internality for success subscale scored significantly higher in subtraction than those not receiving feedback, when items were arranged in ascending order; (b) those receiving feedback and who scored between 0 and .7 (out of 6) on the stable internality for failure subscale scored significantly higher in subtraction than those not receiving feedback, when items were arranged in descending order; (c) those receiving feedback and who scored roughly between 1.0 and 3.4 (out of 6) on the unstable internality for failure subscale scored significantly higher in subtraction than those not receiving feedback, when items were arranged in descending order; and (d) those receiving feedback and who scored roughly between 3.9 and 8.0 (out of 12) on externality for success or failure subscale scored significantly higher in subtraction than those not receiving feedback, when items were arranged in ascending or descending order.

Discussion

Locus of control does explain to some degree the effect of response feedback upon diagnostic mathematics test performance. Superior subtraction performance appeared for individuals who tend to attribute their achievement success to internal causes (ability or effort) or their failure to unstable internal causes (lack of effort), at least when test items were ordered in an ascending or descending fashion so as to emphasize the effect of feedback.

Although it was hypothesized that when feedback suggesting failure was maximized under a descending schedule it would be debilitative for those tending to attribute their failure to stable internal causes (lack of ability), this was not the case; no difference in performance was found. However, subjects below the mean stable internal locus benefited from feedback. One might interpret this as a kind of back-handed support for the hypothesis, but no theoretical explanation seems readily apparent.

It was puzzling to find that individuals who tended to attribute their success or failure to external causes (task difficulty or luck) -- i.e., above the mean -- performed better in subtraction when receiving feedback, regardless of the item arrangement. Somehow, contrary to the hypothesis, it appears that such individuals found both positive and negative feedback as facilitating. We have no reasonable explanation for this result.

The locus measure, as designed for this investigation, appeared to be a useful instrument. It is a theory-based measure (a la Weiner), lending it some construct validity, and the support we found for the two hypotheses add some empirical validation. No one responded randomly, and most had near identical response patterns on the general and mathematics achievement parts of the measure.

The present investigation, while offering some potentially helpful clarifications of the impact of item-by-item knowledge of response feedback on students' test performance, raises several questions. Why were no differences found for performance on the addition subscale? Had students so well mastered addition by the third grade that an effect could not be measured? This was probably the case, since most subjects' performance deteriorated as they encountered items in addition of fractions/ decimals. Did the favorable (though not significantly so) performance on the descending item arrangements reflect a scoring procedure that yields an overly optimistic measure of students' performance or do students generally perform better when items are arranged in a descending fashion? Reflection on the testing suggests the former is probably true. In the ascending version, students tended to miss an occasional item before reaching a plateau, while in the descending version there seemed to be a tendency to miss all items until reaching an appropriate level. Since those items sporadically missed in the ascending arrangement would be assumed correct in the descending arrangement, this would result in higher scores for the descending item arrangement. This highlights the inadvisability of using only total performances in a diagnostic testing situation. Finally, since to some degree all hypothesis testing revealed significantly higher performance under feedback, could the locus of control theorizing be replaced by a simple arousal hypothesis? That is, regardless of its nature, feedback is arousing and hence leads to heightened performance. Although support for this might be found in Table 4, the differential effect found between those high and low in the attributions studies argues against this.

In light of these questions and the only partial support for the theorizing presented in this study, further investigations are needed. The present investigation continued to support the benefits of item-by-item knowledge of response feedback in the diagnostic mathematics test situation. However, continued investigation into the explanation for that superiority is needed. It may be, for example, that ability level, achievement anxiety or arousal are constructs which, when considered concurrently with locus of control, may more fully explain the effects of feedback. Gilmer (Note 1) has argued a case for relating feedback effects and anxiety.

Until a complete explanation is found for the role of feedback in diagnostic testing, its consistent and widespread advocacy should not be made. Yet, the evidence to date suggests that educators should probably not fear adverse results in responding when a child requests correct - incorrect feedback on a diagnostic test item.

REFERENCES NOTES

1. Gilmer, J.S. *The effects of immediate feedback versus traditional no-feedback in a testing situation.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, 1979.

2. Hambleton, R.K. *The effects of item order and anxiety on test performance and stress.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 1968.

3. Towle, J.J. & Merrill, P.F. *Effects of anxiety and item difficulty sequencing on mathematics aptitude test performance.* (Technical Memo Number 46). Tallahassee, FL: Florida State University, Computer-Assisted Instruction Center, 1972.

4. Zeineh, F.A., Pierce, E. & Silverman, M. The effect of test item ordering on the performance of children of different test anxiety levels. In Romberg, T.A. (Ed.) *Four studies on aspects of assessing computational performance.* (Technical Report No. 297). Madison, WI: University of Wisconsin, Wisconsin Research and Development Center for Cognitive Learning, 1975.

5. Butkowsky, I.S. *General versus task-specific measures of children's causal attributions: Their relationship to achievement behavior and reading ability.* Paper presented at the Annual Meeting of the American Educational Research Association, Boston, 1980.

6. Ceurvorst, R.W. *Johnny: A program to perform analyses using the Johnson-Neyman technique.* (Technical Report Number 14) Tempe, AZ: Arizona State University, University Testing Service, 1978.

REFERENCES

Anastasi, A. *Psychological testing* (4th ed.) New York: Macmillan, 1976.

Bar-Tal, D. Attributional analysis of achievement-related behavior. *Review of Educational Research,* 1978, 48, 259-271.

Betz, N.E. Effects of immediate knowledge of results and adaptive testing on ability test performance. *Applied Psychological Measurement,* 1977, 1, 259-266.

Brenner, M. The effect of item-difficulty on total test reliability, difficulty, and discrimination. Journal of Applied Psychology, 1964, 48, 98-100.

Connolly ,A.J., Nachtman, W. & Pritchett, E.M. Key Math Diagnostic Arithmetic Test. Circle Pines, Minnesota: American Guidance Service, 1971.

Crandall, V.C., Katkovsky, W. & Crandall, V.J. Children's beliefs in their control of reinforcements in intellectual academic achievement behaviors. Child Development, 1965, 36, 91-109.

Engelhardt, J.M. Diagnosis and remediation in school mathematics: Developing continuity among R and D efforts. In Heddens, J.W. & Aquila, F.D. (Eds.) Proceedngs of the third national conference on remedial mathematics. Kent, Ohio: Kent State University, 1976.

Engelhardt, J.M. Feedback in diagnostic testing. In Speer, W.R. (Ed.) Clinical investigations in mathematics education. Bowling Green, Ohio: Research Council on Diagnostic and Prescriptive Mathematics, 1978.

Klosner, N.C. & Gellman, E.K. The effect of item arrangement on classroom test performance: Implications for content validity. Educational and Psychological Measurement, 1973, 33, 413-418.

Lefcourt, H.M. Locus of control: Current trends in theory and research. New York: John Wiley and Sons, 1976.

Marso, R.N. Test item arrangement, testing time and performance. Journal of Educational Measurement, 1970, 7, 113-118.

Rust, J.O., Strang, H.R., and Bridgeman, B. How knowledge of results and goal setting function during academic tests. Journal of Experimental Education, 1977, 45, 52-55.

Sax, G. & Carr, A. An investigation of response sets on altered parallel forms. Educational and Psychological Measurement, 1962, 2, 371-76.

Sax, G. & Cromack, T. The effects of various forms of item arrangements on test performance. Journal of Educational Measurement, 1966, 3, 309-311.

Shoemaker, D.M. Test statistics as a function of item arrangement. Journal of Experimental Education, 1970, 39, 85-88.

Weiner, B., Heckhausen, H., Meyer, W., & Cook, R.E. Causal ascriptions and achievement behavior: The conceptual analysis of effort. Journal of Personality and Social Psychology, 1972, 21, 239-248.

Figure 1. Attributions represented by response choices.
(from Weiner, Heckhausen, Meyer and Cook, 1972)

		Locus of Control	
		Internal	External
Stability	Stable	ability	task difficulty
	Unstable	effort	luck

Figure 2. Possible scores from locus of control measure.

Locus	Stability	Feedback	(Max Score)	Achievement Situation
Internal	Stable	Success	(3)	General Mathematics or Combined (double the max score shown)
		Failure	(3)*	
		Combined	(6)	
	Unstable	Success	(3)	
		Failure	(3)*	
		Combined	(6)	
	Combined	Success	(3)*	
		Failure	(3)	
		Combined	(6)	
External	Stable	Success	(3)	
		Failure	(3)	
		Combined	(6)	
	Unstable	Success	(3)	
		Failure	(3)	
		Combined	(6)	
	Combined	Success	(3)	
		Failure	(3)	
		Combined	(6)*	

Table 1

Means (Standard Deviations) for Locus of Control including a t-test Comparing Subscales (n=36)

Item	External	Internal						
		Success			Failure			Total
	Total M (SD)	Stable M (SD)	Unstable M (SD)	Combined M (SD)	Stable M (SD)	Unstable M (SD)	Combined M (SD)	M (SD)
General Ach. Subscale	2.28 (1.43)	.61 (.80)	2.00 (1.01)	2.50 (.78)	.17 (.45)	.89 (1.01)	1.22 (1.02)	3.72 (1.43)
Math Ach. Subscale	2.25 (1.50)	.53 (.56)	2.11 (.92)	2.75 (.65)	.39 (.77)	.83 (.88)	1.17 (1.03)	3.92 (1.13)
Total	4.47 (2.35)	1.17 (1.06)	4.17 (1.73)	5.22 (1.20)	.56 (1.03)	1.69 (1.65)	2.42 (1.87)	7.64 (2.34)
t	.14	.55	-.94	-1.86	-1.85	.36	.40	-1.02
p	.89	.59	.35	.07	.07	.72	.69	.31

Table 2

Means (Standard Deviations) for Locus of Control and One-Way ANOVA Comparing Feedback and No-Feedback Groups.

Item/Group	n	External	Internal						
			Success			Failure			Total
		Total M (SD)	Stable M (SD)	Unstable M (SD)	Combined M (SD)	Stable M (SD)	Unstable M (SD)	Combined M (SD)	M (SD)
Possible Score		12	6	6	6	6	6	6	12
Feedback	18	4.39 (2.57)	.94 (1.11)	4.17 (2.12)	5.06 (1.35)	.67 (1.19)	1.78 (1.63)	2.56 (1.88)	7.61 (2.57)
No Feedback	18	4.56 (2.18)	1.39 (.98)	4.17 (1.29)	5.39 (1.04)	.44 (.86)	1.61 (1.72)	2.28 (1.90)	7.67 (2.17)
F (1/34 df)		.04	1.62	.00	.69	.42	.09	.19	.00
p		.99	.21	.99	.99	.99	.99	.99	.99

Table 3

Diagnostic Mathematics Test Performance Means and Standards Deviations for Group 1 (Ascending - Descending) and Group 2 (Descending - Ascending), including *t*-Tests for Differences Between Groups and Between Versions. Addition scale 15 items, subtraction scale 14 items.

Group/Scale	Ascending Version				Descending Version				Ascending vs. Descending Difference			
	Group 1	Group 2			Group 1	Group 2			Group 1		Group 2	
	M(SD)	M(SD)	*t*	*p*	M(SD)	M(SD)	*t*	*p*	*t*	*p*	*t*	*p*
Feedback Group (n)	9	9			9	9						
Addition	10.56 (1.24)	10.00 (1.87)	.74	.47	11.00 (.87)	10.11 (1.54)	1.51	.15	1.08	.31	.43	.68
Subtraction	9.00 (1.00)	8.00 (2.06)	1.31	.21	9.33 (1.00)	8.33 (1.58)	1.60	.13	1.15	.28	1.41	.20
No Feedback (n)	9	9			9	9						
Addition	10.11 (1.36)	9.89 (2.09)	.27	.79	11.11 (1.17)	10.11 (1.54)	1.55	.14	1.46	.18	.41	.70
Subtraction	8.00 (1.66)	6.44 (.88)	2.48	.02	8.33 (1.80)	7.00 (.71)	2.07	.07	.82	.44	1.89	.10
Combined Groups (n)	18	18			18	18						
Addition	10.33 (1.28)	9.94 (1.92)	.71	.48	11.06 (1.00)	10.11 (1.49)	2.23	.03	1.83	.08	.57	.58
Subtraction	8.50 (1.42)	7.22 (1.73)	2.42	.02	8.83 (1.50)	7.67 (1.37)	2.43	.02	1.37	.19	2.41	.03

Table 4

Comparison of Mathematics Performance for Those Receiving Feedback and Those Not Receiving Feedback

Item/Group	n	Addition		Subtraction	
		Ascending M(SD)	Descending M(SD)	Ascending M(SD)	Descending M(SD)
Feedback	18	10.28 (1.56)	10.56 (1.29)	8.50 (1.65)	8.83 (1.38)
No Feedback	18	10.00 (1.72)	10.61 (1.42)	7.22 (1.52)	7.67 (1.50)
t		.51	-.12	2.42	2.43
p		.62	.90	.02	.02

Table 5

Regression of Locus of Control on Mathematics Test Performance (df=1/32)

Scale/Group		External				Internal											
						Success				Failure							
		Total				Combined				Stable				Unstable			
	n	r	slope	intercept	F (p)	r	slope	intercept	F (p)	r	slope	intercept	F (p)	r	slope	intercept	F (p)
Addition (Ascending)																	
Feedback	18	-.47	-.28	11.53	2.89 (.10)	.16	.18	9.34	.50 (.49)								
No Feedback	18	.14	.11	9.49		-.10	-.16	10.88									
Addition (Descending)																	
Feedback	18	-.26	-.13	11.14	.08 (.78)					.20	.22	10.41	.05 (.82)	.31	.25	10.11	.43 (.52)
No Feedback	18	-.29	-.19	11.47						.20	.33	10.46		.08	.06	10.51	
Subtraction (Ascending)																	
Feedback	18	.14	.09	8.09	1.21 (.28)	-.17	-.21	9.56	.52 (.48)								
No Feedback	18	-.24	-.16	7.97		.09	.13	6.50									
Subtraction (Descending)																	
Feedback	18	.02	.01	8.79	3.62 (.07)					-.14	-.17	8.94	1.95 (.17)	.24	.21	8.46	.34 (.57)
No Feedback	18	-.54	-.37	9.35						.31	.54	7.43		.04	.03	7.61	

Table 6

Johnson-Neyman Compariosn of Regression Lines for Regions of Significant Difference Along the Covariate Axis Between Feedback and No Feedback Groups (Locus of Control Regressed on Mathematics Test Performance)

Mathematics Scale	External		Internal					
			Success		Failure			
	Total		Combined		Stable		Unstable	
	Intersection	Region (Actual Score Range)	Intersection	Region (Actual Score Range)	Intersection	Region (Actual Score Range)	Intersection	Region (Actual Score Range)
Addition (Ascending)	5.14	None	4.42	None				
Addition (Descending)	5.96	None			-.53	None	2.14	None
Subtraction (Ascending)	-.46	3.90 to 9.14 (0 to 8.00)	8.90	3.93 to 5.56 (2.00 to 6.00)				
Subtraction (Descending)	1.47	3.98 to 46.29 (0 to 8.00)			2.16	-2.60 to .73 (0 to 4.00)	-4.90	1.14 to 3.41 (0 to 6.00)

Review of Research on the Validation of Learning Hierarchies

John D. McGregor, Ph.D.
Christopher Newport College

In a series of studies dating from the early 1960's, Gagne and his colleagues (Gagne and Paradise, 1961; Gagne, 1962; Gagne et al, 1965; Okey and Gagne, 1970) have examined the hypothesis that hierarchical sequencing of certain learning outcomes (i.e., intellectual skills) maximizes positive transfer. This hypothesis states that learning the skills that come early in the sequence is necessary for learning later skills. Gagné termed the formal structure arising from this sequencing a learning hierarchy (Gagné, 1962). One review of learning hierarchy research (Walbesser and Eisenberg, 1972) summarized the substantial research that supports the hypothesis. Several studies have clarified many of the characteristics of learning hierarchies (Gagné, 1968a; Gagné, 1970; Okey, 1973) and others have also applied learning hierarchies to a wide range of subject matter (Beeson, 1977; Okey and Gagné, 1970; Riban, 1971; Treffinger and Huber, 1975; Wang et al, 1971; Gagné, 1970).

Subsequent research has branched into several issues. Many researchers have considered the use of learning hierarchies in planning classroom instruction. Gagné was not certain "that a learning hierarchy is supposed to represent a presentation sequence for instruction in an entirely uncomplicated way". (Gagné, 1968b, p. 4). Others however, have not been as hesitant about the relationship between hierarchies and content sequencing. Okey (1973, p. 88) stated that "a hierarchy of tasks . . . can serve as an instructional map or guide for a teacher because it is a list of which tasks are to be accomplished by students as well as the sequence in which they should be studied and taught".

Hierarchical sequencing has also been incorporated into instructional design systems. Briggs (1970) includes in his system a step in which terminal outcomes are analyzed to establish a sequence of objectives. Bloom (1956), Gagné (1970) and Merrill (1971) all describe techniques for sequencing objectives into hierarchies, and Tennyson and Merrill (1971) compared the three systems.

Learning hierarchies have also been used in research on the learning process. Using learning hierarchies, Weigand (1970) studied intellectual development in children. Trembath and White (1975) discussed the use of hierarchies in promoting mastery learning. Gagné (1968a) discussed the implications of a cumulative learning model for transfer of learning.

Finally, hierarchical sequencing has been used by test designers. Skills to be tested are arranged in a hierarchy. Subjects passing a test at a given level can be assumed to be capable of passing tests on any skill lower in the hierarchy (Resnick and Wang, 1969).

This variety of uses has inevitably led to different definitions of a valid hierarchy and to different concepts of how hierarchies can be developed and validated. In a review of the literature, several elements of the validation process require distinction:

1. Some validation techniques judge validity of a fixed set of hierarchical dependencies while others restructure the set of dependencies to maximize some variable.

2. Some statistical measures used for validation are based upon standard statistical procedures and distributions while other measures are indices developed specifically for hierarchy validation and thus lack the stable interpretability of traditional statistical concepts.

3. Some researchers test for a skill after the entire instructional sequence while others include the testing within the sequence.

4. Some validation techniques include measures that examine the learning sequence as well as the hierarchy while others also review the validity of the test questions used to gather data.

5. Some researchers gather data from students who have been given an instructional sequence based on the hierarchy while others gather data from test questions alone.

The purpose of this paper is to examine five of the most widely used learning hierarchy validation models, including Gagné's, White's and Walbesser's models and the Intraconcept Analysis and Multiple Scalogram Analysis models. In the review of the models, problems common to many of the models will be identified. Possible solutions to these problems will be identified and research needed to verify these solutions will be suggested. Whether these models are variations on a single construct or are significantly different will be considered.

Definition of Validity

Gagne initially hypothesized that a valid hierarchy was one whose elements were arranged to maximize upward positive transfer. In several studies (Gagné & Paradise, 1961; Gagné et al, 1962) Gagné utilized an index (i.e., proportion positive transfer) to measure agreement with the hierarchical hypothesis. After the skills were taught in the hypothesized order, Gagné calculated the number of pass-fail patterns consistent with the hypothesis of positive transfer.

A valid hierarchy, according to Okey (1973, p. 96), is one which "describes the actual learning sequence of a high percentage of learners". Kane, McDaniel and Phillips (1971) define a valid hierarchy as one which yields an optimal learning sequence. Phillips (1971) defines an optimal learning sequence as one which maximally facilitates achievement, transfer and retention; requires the least amount of time for the learner to complete; and induces minimal anxiety and frustration on the learner's part. Niedermeyer, Brown and Sulzen (1969) and Brown (1970) found that subjects using materials that were logically sequenced and based on learning hierarchies took less time to complete the instructional program than subjects using randomly sequenced materials. Differences in favor of the logically sequenced group were also found on the number of errors on the criterion test of problem-solving skills.

According to Resnick and Wang (1969), designers of educational tests view a valid hierarchy as an ordering in which students passing a test on a top-level objective in the hierarchy could also pass tests on all objectives lower in the hierarchy. They see this requirement as closely related to, but not the same as positive upward transfer. "Researchers interested in the use of hierarchies as a means of sequencing instructional objectives, therefore, are necessarily concerned that hierarchy validation studies seek to establish independently the scaling properties of hierarchies and their learning transfer properties" (Resnick & Wang, 1969, p. 4). Airasian (1971, p. 54) describes a valid hierarchy as one in which "(1) mastery of a task at any sequential level is indicative of mastery of all earlier sequential tasks and (2) failure on any task is predictive of failure on all later tasks".

Interest in developing a sequence of skills leading to a terminal skill led to the definition of a valid hierarchy as one which promotes upward positive transfer. Other investigators including Airasian have added the condition that passing a skill implies the passing of all prerequisite skills. Resnick and Wang (1969) noted that these are significantly different requirements. They also pointed out that the first definition should be used by those interested in developing instructional sequences. Studies using both definitions are included in the models reviewed below.

Gagné's Model

Acceptance of a particular definition of validity as well as the intended use for the hierarchy influence the design of the validation study. Gagné's initial research established a model used by Gagne and others in several studies (Gagné, 1962; Gagné & Paradise, 1961; Gagné et al, 1962; Gagné et al, 1965; Gagné & Bassler, 1963; Kolb, 1967-1968). The model can be summarized as follows:

1. Select a terminal task.

2. Use task analysis to derive subordinate tasks.

3. Construct an instructional program to teach elements of

the hierarchy.

4. Have a group of subjects use the instructional program.

5. Test achievement of elements of the hierarchy by subjects.

6. For each connection construct a 2x2 table for pass-fail combinations on both the lower and upper skills.

7. Remove connections from the hierarchy where a large percentage of students were able to learn the upper skill without learning the lower skill.

Gagné used a measure, proportion positive transfer, as his measure of validity. Subjects were divided into four groups:

(1) number of Ss who answered correctly the question for the higher skill and the question for the lower skill;

(2) number of Ss who answered incorrectly the question for the higher skill and the question for the lower skill;

(3) number of Ss who answered correctly the question for the higher skill and incorrectly the question for the lower skill;

(4) number of Ss who answered incorrectly the question for the higher skill and correctly the question for the lower skill.

The measure was defined as: $\frac{(1)+(2)}{(1)+(2)+(3)}$

Thus the ratio was of those subjects who behaved in a manner consistent with positive transfer to the sum of those who behaved consistently with the theory and those who directly contradicted it. Values of .90 or higher were considered evidence of validity. White (1974b) noted that this index could reach levels near 1 with only a small number of students learning any of the skills in the connection.

In addition, no data were gathered that would indicate whether failure to learn a skill was due to lack of a prerequisite skill or due to inadequate instruction. Subjects in group (4), the most likely victims of poor instruction, were excluded from any calculation. Okey (1973) noted that either fault could lead to low student achievement. White (1974a) listed several criticisms of the above model. Of importance here are the following:

1. Hierarchies need a common sense validity. Validating a connection because no one could answer the upper questions without answering the lower questions is a necessary but not sufficient condition for a valid connection.

2. Some studies have used too few subjects. If some subjects can learn the skills out of order, then in a small sample their number may be under-represented.

3. Several studies using one test question per element could give no reliability estimate for the achievement test.

4. Delaying achievement testing to the end of the instructional program can lead to incorrect decisions. Gagné and Bassler (1963) showed that forgetting can occur in lower hierarchical elements, leading to the rejection of otherwise valid connections.

5. Critical values of the statistics used to judge the validity of connections in the hierarchy were arbitrarily set.

White's Model

White (1974a) modified Gagné's model in several ways. Of most importance are the following:

1. He checked the reasonableness of the hierarchy with experienced subject-matter experts.

2. He tested whether elements of the hierarchy represent single-dimension skills. White (1971) described the procedures to be followed in such a study.

3. He included the achievement test questions in the instructional program. Immediate testing gives an accurate picture of whether the skill has been learned.

4. He used a statistical test of hierarchical relationship based on the binomial distribution. White and Clark (1973) described such a test.

White (1971, 1974c) used this model to investigate a hierarchy of graphical skills. Beeson (1944) and Linke (1973, 1974, 1975) also used modified versions of this model.

The statistical test developed by White and Clark is based on the binomial distribution. Unlike other hierarchy tests, this test does not require that a pass-fail decision be made when more than one question is used to measure achievement of a skill. Even though the test includes compensation for errors of measurement, the studies using this model have failed to report any measure of reliability of the test questions. While this statistical test represents an improvement on Gagné's proportion positive transfer, the test does not differentiate between missing tasks and poor instruction. In determining whether a connection is valid or not, the test uses only a small proportion of the total data for the connection. For example, where a two question achievement test is used a 3x3 table (0,1,2 questions correct) is formed. The test only examines the cell where both upper skill questions were wrong. Other cells which also contain data which could contradict the hierarchy hypothesis (e.g., 1 upper skill question correct, no lower skill questions correct) are not included.

Walbesser's Model

In a modification of Gagné's model, Eisenberg and Walbesser (1971) developed a series of five ratios based on Gagné's four groupings of response patterns. Eisenberg (1970) and McKeen (1970) reported studies which used versions of this model. Several studies that employed this procedure, including Intraconcept Analysis model, used only three of the ratios. The first of these is consistent with the hierarchical hypothesis. The adequacy ratio measures the degree to which subjects who learned the lower level skills were able to attain the upper level skill. The completeness ratio measures the percentage of subjects who learned any of the skills in the connection. With the value of each ratio ranging from zero to one, Walbesser established 0.85 as the critical value for each of the ratios. However, as with Gagné's proportion positive transfer, there is no statistically valid procedure for determining these critical values. Phillips (1971) determined that when instructional sequences are not included in the validation process, the acceptable levels are (a) consistency ratio 0.85, (b) adequacy ratio 0.70, and (c) completeness ratio 0.50.

The three ratios comprehensively measure the validity of a connection, particularly when a learning sequence is used as part of the validation study. McGregor (1976) demonstrated that Walbesser's procedure did not give results similar to White's technique when a large proportion of students did not do well on the instruction for a connection. Post hoc comparisons of Walbesser's consistency ratio and White's statistical test showed a significant amount of agreement. The adequacy and completeness ratios failed to reach critical values not because the skills were not arranged in a hierarchical manner but rather because additional skills or improved instruction were needed for an effective instructional sequence.

Models reviewed to this point have hypothesized a set pattern of hierarchical relations among tasks and have sought to validate or reject the relations. Some models alter the hypothesized hierarchical relationships on the basis of student performance data. Measures such as item difficulty (Causey 1976) or Walbesser's ratios (Uprichard and Phillips, 1977) have been used to restructure the hierarchical connections. Intraconcept analysis and Multiple Scalogram Analysis are two methods that use performance data to develop hierarchy connections.

Intraconcept Analysis

Intraconcept analysis utilizes a modification of Gagné's task analysis. According to Uprichard and Phillips (1977), in addition to considering content factors, intraconcpet analysis also considers psychological factors in developing and ordering the tasks. While Gagné's task analysis begins at the terminal objective and moves down, creating each level, intraconcept analysis creates a set of classes of tasks and orders the classes. Each class is then laterally divided into a set of tasks that Uprichard and Phillips termed sum categories, and questions are then generated to measure attainment of each skill. Most studies using intraconcept analysis have used indirect validation. After tests are administered to a wide range of subjects, the data from these tests is used to validate the hierarchical connections.

Various measures have been used to validate hierarchies in intraconcept studies. Causey (1976) initially ordered tasks by item difficulty. Hierarchies, first linear and then branching were hypothesized and tested using Walbesser's ratios. But Causey (1978) cautioned that studies conducted in this manner were actually generating hierarchies which should then be validated using a separate population. Hill (1976) performed a similar study.

Uprichard and Phillips (1976, 1977) in several intraconcept studies have used a combination of Walbesser's ratios and pattern analysis. Using indirect data, Walbesser's ratios were calculated for each connection in the hypothesized hierarchy. The index of agreement for the pattern analysis was calculated for the entire hierarchy. This index measures agreement between observed patterns of response and the expected pattern. The expected pattern is that after a student fails to achieve a task, all higher tasks will also be failed. In one study (Uprichard and Phillips, 1977), tasks were rearranged to maximize the Walbesser ratios, but no arrangement reached criteria for all 3 ratios on all tasks. For the hierarchy with maximum Walbesser ratios, an index of agreement of 0.70 was calculated. Since no test of significance has been developed for this index, no definite conclusion can be reached. However, it implies that a large percentage of subjects varied from the expected pattern on at least one connection.

An important improvement in this study was the reporting of a measure of internal consistency for the test used to gather data. Another improvement is the combination of a statistic which measures the validity of a single connection and a statistic which measure the validity of the entire hierarchy. The intraconcept model does not include an investigation of possible further division of the tasks, though White (1971) cited the need for such an investigation and outlined procedures for such a study. In Uprichard and Phillips (1977) 89 tasks were combined to make 45 tasks. If these resulting tasks were not singular skills, the test questions may not have measured all facets of a task. In his criticisms of Gagné's model, White cited the need for a common-sense validity for the arrangement of elements in the hierarchy (1974a). This is a particular concern when elements of the hierarchy will be statistically arranged. Although the tasks in an intraconcept study are logically derived, the final order of elements is statistically determined.

Multiple Scalogram Analysis Model

Scalogram analysis was developed by Guttman (1944). Guttman's coefficient of reproducibility gave a measure of the degree to which items form a perfect scale. In a perfect scale, the ability to perform a task implies the ability to perform all tasks that are placed lower in the scale. In scalogram analysis, tasks are rearranged to maximize the coefficient of reproducibility. D'Mello and Willemsem (1969) reported a study in which four tasks on beginning number concepts were hierarchically arranged. Scalogram analysis was used to restructure the original order. The final ordering had a coefficient of reproducibility of 0.99. Scalogram analysis constructs a single linear scale from a set of tasks, a form which is inadequate for all but the simplest hierarchy. Lingoes (1963) developed a modification term multiple scalogram analysis which generates multiple linear scales.

Resnick and Wang (1969), Resnick, Wang and Kaplan (1970) and Wang, Resnick and Boozer (1971) reported results of validation studies using multiple scalogram analysis. Resnick (1967) described a process of behavior analysis that derives from a given task a series of tasks and orders these tasks into a hierarchy. Using this procedure, Wang, Resnick and Boozer (1971) developed a series of hierarchies for some early mathematics behaviors. Several hierarchies were supported by the data. Resnick & Wang (1969) reported development of hierarchies of counting behaviors. Although many of the hypothesized hierarchies were rearranged, coefficients of reproducibility higher than 0.9 were reported for 5 scales. Much of the research using this model deals with predicting lower-skill behavior from higher-skill performance. Resnick and Wang (1969) distinguished between "prediction downward" and "learning facilitation upward". While these concepts are closely related, they are not identical. Resnick and Wang (1969) discussed complications in the use of Multiple Scalogram Analysis. The linear scales generated by this analysis must be independent in the sense that once a skill is selected for one scale, it cannot be considered for any other scale. Thus, to validate an entire hierarchy, each possible linear path must be separately tested. It was also noted that because of basic assumptions about the model, testing a scale of two skills is a questionable practice. This implies that unlike the other models reviewed here, multiple scalogram analysis cannot be used to validate individual connections within a hierarchy.

This model has no statistical sampling distribution and does not account for errors of measurement (White, 1974b). White also pointed out that Resnick and Wang (1969) found that when using a complex branching hierarchy, the resulting set of scales was difficult to interpret. Hofman (1977) stated that "the Guttman model assumes that the skills . . . are ordered according to their difficulty". Item difficulty alone, however, does not adequately determine hierarchical orderings (Phillips, 1971).

Although the five models reviewed here represent the major approaches to validation, other measures have been used. An ordering theoretic method developed by Airasian has been used in several studies (Airasian, 1971; Airasian & Bart, 1975; Bart & Airasian, 1974; Bart & Krus, 1973). Using student performance data, this method examines 2x2 contingency tables to calculate an index, "conditional item difficulty," for all possible combinations of hierarchy elements. While similar to Multiple Scalogram Analysis, no a priori assumptions about relationships are necessary. Ordering-theoretic analyses use a preset tolerance level since the method does not deal with the occurrance of random error in student response data (Airasian & Bart, 1975). Ordering-theoretic methods also lack a statistical sampling distribution as a basis.

Validation of a learning hierarchy is a complex process involving many steps. Except for multiple studies by the same investigator, comparisons of studies using a particular model indicate procedural differences. Whether these differences are significant or slight, any statement about studies using a model can be taken only as a generalization. This review summarizes the major approches to validation and serves as a basis for the discussion that follows.

Discussion

Gagné's original model and several others (Okey & Gagne, 1970; White, 1971; White, 1974c) required task analysis not only to generate the elements of the hierarchy but to establish the order of relationships between elements as well. Statistics were then used either to test the hypothesized relationships. Gagné's and White's statistics sought to determine if an unacceptably large number of students could learn a task without learning the prerequisites. Recent studies have utilized statistical measures to rearrange elements of the hierarchy and thereby establish different relationships among the elements than those proposed in the task analysis. Multiple Scalogram Analysis (Resnick & Wang, 1969) and ordering-theoretic methods (Airasian, 1971) are methods developed specifically to order a set of objectives. Causey (1976) rearranged elements of a hierarchy according to item difficulty while Uprichard and Phillips (1976) reordered based on values of Walbesser's ratios. The use of these methods raised two questions: (1) What principles determine the order of the elements? (2) Do the statistics justify that order?

Gagne (1974) has indicated two types of analysis that decompose tasks into hierarchical sequences. He (1965) proposed a hierarchy of learning types running from simple to complex tasks and hypothesized that any intellectual skill could be broken into a hierarchy of prerequisite learning types such as rules, concepts, discrimination, and so on. Soulsby (1975) has noted several criticisms of this notion. Similar hierarchical classification schemes have been proposed by Bloom (1956) and Merrill (1971).

The second type of hierarchical sequencing has been termed a logical hierarchy (Tennyson & Boutwell, 1971). In this case the structure of the hierarchy is determined by the structure of the subject matter. A terminal objective would be task analyzed to develop a hierarchy based on prerequisite knowledge rather than prerequisite learning type. This type of analysis has dominated recent learning hierarchy studies.

Gagne (1974) has also indicated that a given hierarchy might contain both types of analysis. In would appear, however, that procedures for validating the two types of sequences should be very different. In the first instance, two types of studies are required. First, the hypothesized ordering of learning types should be validated. Such studies have been performed on Bloom's taxonomy (Kropp, Stoker & Bashaw, 1966; Stoker & Kropp, 1964). Following this general study, specific studies of each hierarchy would determine if the objectives were of the appropriate type and if they were in the proposed order. Statistics that rearrange elements based on item difficulty or scalability are particularly unsuited for this type of study.

For logical hierarchies, no general sequencing study is possible. Rather, each hierarchy must be considered in the light of student data. Uprichard and Phillips have indicated that psychological information about the learners for whom the hierarchy is being developed can be very helpful in the original sequencing (Uprichard and Phillips, 1977). To conform to Gagné's original definition of hierarchical learning, elements must be arranged so that learning a lower-level skill should facilitate learning the upper level skill. White (1974a) presented a rigorous model for validation using this hypothesis. While White acknowledges that the time and number of subjects required to test such a model make this best possible design impractical, the models reviewed earlier in this article represent the attempts of various investigators to approximate the ideal design. Each of these designs requires student performance data. Gagne's original research, many of White's studies and those of several other investigators have used student data (termed "direct data") as they encountered instruction on a topic. Other studies have used indirect data. That is, the data was gathered using only a test instrument designed to measure knowledge of each of the hierarchy elements. Phillips (1974) states a case for the use of indirect data, arguing that one major concern in the use of direct data is the role of forgetting. Gagné and Bassler (1963) found that while learning was hierarchical, forgetting was not. White and Gagné (1978) reported that there is a clear implication that hierarchies of learning and remembering are identical. A definitive study on this question would greatly encourage (or discourage) the use of indirect data.

Indirect data are an approximate measure of content knowledge. Hierarchy validation calls for a measure of the sequence of learning. White (1974a) included a step in his validation model in which a learning program is used with test questions embedded in the program following instruction on each element. Whether this type of control is exercised or not, the validation measures must be concerned with the learning program is used; however, it is presumed the subjects have learned some of the skills included in the hierarchy. Most studies use populations who would have encountered similar instruction on given topics, but very few validation techniques include measures that can point to problems of instruction. Okey (1973) states that "no clear distinction can be made between an inadequate analysis of a task and inadequate instruction for teaching the tasks in sequence . . . Some of the validation procedures . . . can be used to determine the location of a problem not the cause". If the product of usable instructional sequences and not a simple ordering of tasks is desired, this is an important distinction. Walbesser's adequacy ratio can be used to locate a problem connection.

The question of sequence of learning leads one to question the use of validation measures that resequence elements of the hierarchy using performance data. Since no statistic is available to reorder elements on the basis of upward positive transfer, approximations have been used. Multiple Scalogram Analysis, ordering theoretic methods and some intraconcept studies have used some form of item difficulty measure, though item difficulty alone is not a sufficient basis for reordering (Phillips, 1971). Multiple Scalogram Analysis is a less-than-adequate technique since it is theoretically possible for objectives to scale perfectly but for instruction in a task lower in the scale not to produce significant amounts of transfer to higher-level objectives (Resnick and Wang, 1969). Hierarchy connections which have been generated statistically represent post hoc hypotheses and should not be validated using the same set of data. Resnick and Wang (1969) and Causey (1978) both have noted the need for a second investigation to validate the reconstructed hierarchy. In fact, Cotton, Gallagher and Marshall (1977) have called for a second investigation whenever a large portion of the connections of a hierarchy are invalid.

Finally, one must consider the difference between evaluation on a connection-by connection basis and evaluation of the entire hierarchy at once. While Gagné's, White's, and Walbesser's satistics give a value for each connection, pattern analysis (Rinoldi and Grib, 1960) and to a certain extent Multiple Scalogram Analysis give values for a sequence of objectives. White (1974b) has expressed concern that in using scalogram analysis, a single incorrect connection could lead to rejection of the entire hierarchy. Such measures certainly do not provide information detailed enough for locating trouble spots in the hierarchy. On the other hand, it is possible to construct a set of data in which the data for every connection in the hierarchy are sufficient to validate all connections but for which every subject violates the hierarchical ordering at least once. A series of valid hierarchy connections does not insure that individual subjects learn in a hierarchical sequence. While Uprichard and Phillips (1977) used both Walbesser's ratios and pattern analysis to validate the hypothesized hierarchy, no procedure links statistically the two measures. Nor have significance tests been developed for either measure.

Summary and Conclusions

Numerous research studies have supported the notion of a learning hierarchy and its use in various content areas. Many studies, however, have investigated logically-structured hierarchies as opposed to those sequenced according to conditions of learning. These distinctly different types of hierarchies require different validation criteria and techniques. Studies similar to those for Bloom's taxonomy should investigate Gagne's proposed order of learning types. Individual hierarchies could then be validated be examining types of objectives and their ordering. Elements of hierarchies sequenced in this manner should not be subject to statistical reordering.

Several modifications are needed in the selection and use of statistics to validate logical hierarchies.

1. Poor instruction can lead to an otherwise valid hierarchical connection being rejected. Any statistic used to evaluate student data should also indicate whether the data could have been caused by inadequate instruction. If the hierarchy is to be the basis for an instructional sequence, the statistic should also indicate whether additional intervening tasks would further promote transfer. Walbesser's ratios are the only measures of hierarchical relationship that can give such indications.

2. A hierarchy composed of individually validated connections does not necessarily satisfy the definitions of "optimal learning sequence" or "learning sequence of a large percentage of learners". At the same time, measures which summarize the validity of a hierarchy into a single value tend to average good and bad connections. A combination of techniques is needed to evaluate individual connections and the hierarchy as a whole. Pattern analysis is a promising technique for validating the overall sequence; however, a procedure for statistically linking it to a connection-validation technique such as Walbesser's ratios is needed.

3. With the exception of White and Clark's test for hierarchical dependency, none of the statistics used for validation includes a test of significance. Investigators have claimed validity based on a wide range of values for various procedures. The development of significance tests, particularly for Walbesser's ratios and pattern analysis, would be a useful addition.

Modifications are also needed in the experimental design of validation studies.

1. Some form of task analysis remains the best procedure for

deriving and ordering elements of a hierarchy. The inclusions of psychological considerations in the Intraconcept Analysis model may prove useful; however, a more exacting description of the process is needed to standardize its use.

2. Studies which rearrange hierarchies based on student performance data must be considered as generating rather than validating hierarchies. The newly derived hierarchies should be treated as post hoc hypothesis and subjected to cross-validation.

3. Recent studies have reported reliability measures for the data gathering instruments. Investigators must also be concerned about the validity of these instruments. Validation studies should investigate whether each element of the hierarchy represents a singular task.

4. The use of indirect data assumes that learning of at least some of the tasks in the hierarchy has previously taken place. The sample should be chosen to provide a wide range of original learning sequences. Samples should cross school, district and even state boundaries. Factors such as interference by subsequent instruction should be studied to determine their effect on data gathering. The use of indirect data assumes that if a student forgets prerequisite skills, the terminal skill is also forgotten. Insufficient evidence is available to support or refute this hypothesis. Until such evidence is available hierarchies validated by indirect data should be considered retention hierarchies rather than learning hierarchies.

BIBLIOGRAPHY

Airasian, Peter W. "A Method for Validating Sequential Instructional Hierarchies." Educational Technology, 1971, 11, 54-56.

Airasian, Peter W. & Bart, William M. "Validating a Priori Instructional Hierarchies." Journal of Educational Measurement, 1975, 12, 163-173.

Bart, William M. & Airasian, Peter W. "Determination of the Ordering Theoretic Method." Journal of Educational Psychology, 1974, 66.

Bart, William M. & Krus, David J. "An ordering - Theoretic Method to Determine Hierarchies Among Items." Educational and Psychological Measurement, 1973, 33 291-300.

Beeson, Geoffrey W. "Hierarchical Learning in Electrical Science." Journal of Research in Science Teaching, 1977, 14, 117-127.

Bloom, Benjamin S. (Ed.) Taxonomy of Educational Objectives, Handbook I: Cognitive Domain. New York: David McKay, 1956.

Briggs, Leslie Handbook of Procedures for the Design of Instruction. Pittsburgh: American Institutes of Research, 1970.

Brown, J.L. "Effects of Logical and Scrambled Sequences in Mathematical Materials on Learning with Programmed Instruction Materials." Journal of Educational Psychology, 1970, 61, 41-45.

Capie, W. & Jones, H.L. "An Assessment of Hierarchy Validation Techniques." Journal of Research in Science Teaching, 1971, 8, 137-147.

Caruso, John L. & Resnick, Lauren B. "Task Structure and Transfer in Children's Learning of Double Classification Skills." Child Development, 1972, 43, 1297-1308.

Causey, Matthew. An Empirically Derived Hierarchy of Intraconcept Relationships: Rational Number Multiplication (Doctoral disseration, University of Houston, 1976). University Microfilms No. 77-1503.

Causey, Matthew & Underhill, Robert G. An Empirically Derived Hierarchy of Intraconcept Relationships: Rational Number Multiplication. Toronto, Canada: AERA, 1978. (ERIC Document Reproduction Service No. ED159051)

Cotton, John W., Gallagher, John, P. & Marshall, Sandra P. "The Identification and Decomposition of Hierarchical Tasks." American Educational Research Journal, 1977, 14, 189-212.

Dening, Basil S. "Sequencing Intellectual Skill Development in the Social Studies." The Social Studies, 1976, 67, 63-66.

D'Mello, Sydney & Willemsen, Eleanor. "The Development of the Number Concept: A Scalogram Analysis." Child Development, 1969, 40, 681-688.

Eisenberg, Theodore A. The integration of modified learner-generated sequences into the development of a behaviorally stated learning hierarchy, as applied in mathematics curricula construction. Unpublished doctoral dissertation, University of Maryland, 1970.

Eisenberg, Theodore & Walbesser, Henry H. "Learning Hierarchies - Numerical Considerations." Journal for Research in Mathematics Education, 1971, 2, 244-356.

Gagné, R.M. "The Aquisition of Knowledge." Psychological Review, 1962, 355-365.

Gagné, R.M. Basic Studies of Learning Hierarchies in School Subjects, 1970. (ERIC Document Reproduction Service No.

ED039611).

Gagné, R.M. Conditions of Learning. New York: Holt, Rinehart and Winston, 1965.

Gagné, R.M. "Contributions of Learning to Human Development. Psychological Review, 1968a, 177-191.

Gagné, R.M. "Learning Hierarchies." Educational Psychologist, 1968b, 6, 1-9.

Gagné, R.M. "Learning Hierarchies and Learning Conditions." Journal of Curriculum Studies, 1975, 7, 133-134.

Gagné, R.M. "Some New Views of Learning and Instruction." Phi Delta Kappan, 1970, 468-472.

Gagné, R.M. "Task Analysis - Its Relation to Content Analysis." Educational Psychologist, 1974, 11-18.

Gagné, R.M. & Bassler, Otto, C. "Study of Retention of Some Topics of Elementary Non-Metric Geometry." Journal of Educational Psychology, 1963, 123-131.

Gagné, R.M. & Paradise, N.E. "Abilities and Learning Sets in Knowledge Acquisition." Psychological Monographs, 1961, 75 (14, Whole No. 518).

Gagné, R.M. and Staff of the University of Maryland Mathematics Project. "Some Factors in Learning Non-Metric Geometry." Monographs of Society for Research in Child Development, 1965, 42-49.

Gagné, R.M. et al. "Factors in Acquiring Knowledge of a Mathematical Task." Psychological Monographs, 1962, 76 (7, Whole No. 526).

Guttman, L. A basis for scaling quantitative data. American Sociological Review, IX, 1744, 139-150.

Hill, Tommie Ann An Emperically Derived Hierarchy of Intraconcept Relationships: Whole Number Division (Doctoral Dissertation, University of Houston, 1976). (University Microfilms No. 77-1513).

Hofman, Richard J. "Reproducibility and Learning Hierarchies: A Classification." Journal of Research in Science Teaching, 1977, 14, 563-569.

Hurst, Joe et al. "Hierarchical Analysis of Learning in Economics." Theory and Research in Social Education, 1978, 6, 1-13.

Kane, R.B., McDaniel, E.D. & Phillips, E.R. Analyzing Learning Hierarchies Relative to Transfer Relationships Within Arithmetic. Purdue University, 1971. (ERIC Document

Reproduction Service No. ED053965).

Kolb, J.R. "Effects of Relating Mathematics to Science Instruction on the Acquisition of Quantitative Science Behaviors." *Journal of Research in Science Teaching*, 1967-1968, 174-182.

Kropp, R.P., Stoker, H.W. & Bashaw, W.L. "The Validation of the Taxonomy of Educational Objectives." *The Journal of Experimental Education*, 1966, 34, 69-76.

Lefrancois, Guy "A Treatment Hierarchy for the Acceleration of Conservation of Substance." *Canadian Journal of Psychology*. 1968, 22, 277-284.

Lingoes, James C. "Multiple Scalogram Analysis: A Set-Theoretic Model for Analyzing Dichotomous Items." *Educational and Psychological Measurement*, 1963, 23, 501-524.

Linke, .R.D. The effects of certain personal and situational variables on the acquisition sequence of graphical interpretation skills. Unpublished doctoral dissertation, Monash Univeristy, Clayton, Australia, 1973.

Linke, Russell D. "Influence of Cultural Background on Hierarchical Learning." *Journal of Educational Psychology*, 1974, 66, 911-918.

Linke, R.D. "Replicative Studies in Hierarchical Learning of Graphical Interpretation Skills." *British Journal of Educational Psychology*, 1975, 45, 39-46.

McGregor, John D. A comparison of White's and Walbesser's Learning Hierarchy Validation Techniques Using Addition of Fractions. (Doctoral dissertation, George Peabody College, 1976.) (University Microfilms No. 76-13821)

McKeen, Ronald L. A model for curriculum construction through observations of students solving problems in small instructional groups. Unpublished doctoral dissertation, University of Maryland, 1970.

Merrill, M. David "Content and Instructional Analysis for Cognitive Transfer Tasks. *AV Communications Review*, 1973, 21, 109-125.

Merrill, M. David "Necessary Psychological Conditions for Defining Instructional Outcomes." *Educational Technology*, August, 1971.

Niedermeyer, F.C., Brown, J. & Sulzen, R. "Learning and Varying Sequences of Ninth-Grade Mathematics Materials." *Journal of Experimental Education*, 1969, 37, 61-66.

Okey, James R. "Developing and Validating Learning Hierarchies." *AVCR*, 1973, 21, 87-108.

Okey, James R. & Gagne, R.M. "Revision of a Science Topic Using Evidence of Performance on Subordinate Skills." Journal of Research in Science Teaching, 1970, 7, 321-325.

Phillips, Darrell G. "The Development of the Concept of Displacement Volume: A Hierarchical Model and Its Partial Testing Under Two Methods of Presentation." Journal of Research in Science Teaching.

Phillips, E. Ray. "Development of Optimal Instructional Sequences." Chicago, Illinois: AERA, 1974. (ERIC Document Reproduction Service No. ED097374).

Phillips, E. Ray. "Validating Learning Hierarchies for Sequencing Mathematical Tasks." (Doctoral dissertation, Purdue University, 1971). (University Microfilms No. 72-8008).

Resnick, L.B. Design of a early learning curriculum. Pittsburgh: University of Pittsburgh, Learning Research and Development Centre, December, 1967.

Resnick, Lauren B., ed. "Hierarchies in Children's Learning: A Syposium." Instuctional Science, 1973, 2, 311-362.

Resnick, Lauren B. and Wang, Margaret C. Approaches to the Validation of Learning Hierarchies. Preprint 50, University of Pittsburgh, Learning Research and Development Center, 1969.

Resnick, L., Wang, M. & Kaplan, J. Behavior analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, December, 1970.

Riban, David M. On the Ability to Infer Deficiency in Mathematics from Performance in Physics Using Hierarchies. Journal of Research in Science Teaching, 1971, 8, 67-82.

Rinoldi, H.J.A. & Grib, T.F. Pattern Analysis." The British Journal of Statistical Psycology, 1960, 13, 137-149.

Soulsby, David "Gagne's Hierarchical Theory of Learning: Some Conceptual Difficulties." Journal of Curriculum Studies, 1975, 7, 122-132.

Stoker, H.W. & Kropp, R.P. "Measurement of Cognitive Processes." Journal of Educational Measurement, 1964, 1, 39-42.

Tennyson, Robert D. & Boutwell, Richard C. "A Quality Control Design for Validating Hierarchical Sequencing of Programmed Instruction." NSPI Journal, 1971, 10, 5-10.

Tennyson, Robert D. & Merrill, M. David "Hierarchical Models in the Development of a Theory of Instruction: A Comparison of Bloom, Gagne and Merrill." Educational Technology, 1971, 11, 27-31.

Treffinger, Donald J. & Huber, Jaclyn R. "Designing Instruction in Creative Problem-Solving: Preliminary Objectives and Learning Hierarchies." Creative Behavior, 1975, 9, 260-266.

Trembath, Richard J. & White, Richard T. "Use of Learning Hierarchies in Promoting Mastery Learning." Research in Science Education, 1975, 135-142.

Uprichard, A. Edward & Phillips, E. Ray. "An Intraconcept Analysis of Rational Number Addition: A Validation Study." Journal of Research in Mathematics Education, 1977, 8, 7-16.

Uprichard, A. Edward & Phillips, E. Ray. An Intraconcept Analysis Validation Study. San Francisco, California: AERA, 1976. (ERIC Document Reproduction Service No. ED127142).

Walbesser, H.H. & Eisenberg, T.A. "A Review of Research Behavioral Objectives and Learning Hierarchies." Mathematics Education Reports, ERIC; Columbus, Ohio, 1972.

Wang, Margaret C. "Psychometric Studies in the Validation of an Early Learning Curriculum." Child Development, 1973, 44, 54-60.

Wang, Margaret C., Resnick, Lauren, B. & Boozer, Robert F. "The Sequence of Development of Some Early Mathematics Behaviors." Child Development, 1971, 42, 1767-1778.

Weigand, V. A study of subordinate skills in science problem-solving. In R. Gagne (Ed.) Basic Studies of Learning Hierarchies in School Subjects. Final Report, Project No. 6-2949, U.S. Office of Education, April, 1970.

White, Richard T. "Effects of Guidance, Sequence and Attribute-Treatment Interactions on Learning, Retention and Transfer of Hierarchically Ordered Skills." Instructional Science, 1976, 5, 133-152.

White, Richard T. "Indexes Used in Testing the Validity of Learning Hierarchies." Journal of Research in Science Teaching, 1974b, 11, 61-66.

White, Richard T. Learning Graphical Skills in Kinematics. Unpublished doctoral dissertation. Monash University, Clayton, Australia, 1971.

White, Richard T. "A Model for Validation of Learning Hierarchies." Journal of Research in Science Teaching, 1974a, 11, 1-3.

White, Richard T. "Research into Learning Hierarchies." Review of _Educational Research_, 1973, 43, 361-375.

White, Richard T. "The Validation of a Learning Hierarchy." _American Educational Research Journal_, 1974c, 11, 121-136.

White, R.T. & Clark, R.M. "A Test of Inclusion Which Allows for Errors of Measurement." _Psychometrika_, 1973, 38, 77-86.

White, R.T. & Gagne, R.M. "Formative Evaluation Applied to a Learning Hierarcy." _Contemporary Educational Psychology_, 1978, 3, 87-94.

White, Richard T. & Gagne R.M. "Post and Future Research on Learning Hierarchies." _Educational Psychologist_, 1974, 11, 19-28.

The Development of a Diagnostic Test of Subtraction of Whole Numbers

Ian D. Beattie and Dwight D. Harley
University of British Columbia

It has long been recognized that diagnosis is the key to effective instruction. In a review of thirty one studies dealing with diagnosis and remediation of arithmetic errors Buswell and Judd (1925) were dealing with what are still central elements of the diagnostic process. These are error analysis and the diagnostic interview, both of which are concerned with identification of systematic errors which might reveal a faulty or incomplete grasp of mathematical concepts. Many diagnostic tests date from that time, but not until the decade of the seventies were these elements developed. Present day error analysis is exemplified by the work of Ashlock (1972), Reisman (1972), Cox (1975) and Hopkins (1977). The diagnostic interview has evolved through the administration of a checklist to become integrated into such complex diagnostic models as those of Uprichard (1977) and Speer (1977).

Yet diagnosis has not become an integral part of teaching arithmetic. Theoretically, classroom teachers have a variety of diagnostic tools at their disposal. These include achievement test results, commercial diagnostic tests, checklists, chapter tests in texts, observation and analysis of children's computational work, and oral questioning during instruction. In practice, achievement test results are seldom used for diagnostic purposes, even to identify areas of general concern, diagnostic tests are seldom administered because of their time consuming nature and required expertise, chapter tests are not designed to identify systematic errors, and teacher observation and intuition are not systematic enough to be effective. Also, many systematic errors are not obvious enough, and opportunities for them to be made are not frequent enough, for teacher identification to be feasible. Moreover, the traditional approach to diagnosis relies heavily on individual interviews and requires considerable expertise to be useful. It is therefore impractical.

This test was designed and developed to make it possible for classroom teachers without previous background in diagnosis to make accurate diagnosis of systematic errors. The test is short, normally requiring a maximum of 15-20 minutes to complete and is group administered. In most instances, individual interviews are unnecessary since the test not only identifies the error made but also the faulty concept behind the error, giving direction for future instruction. That is, the test was designed to overcome the major drawbacks of the normal diagnostic process.

In his preface to Mathematics Tests and Reviews, Buros (1975) expressed the view that most standardized tests are of questionable or unknown validity and pretentious in their claims. In the same volume Rogers (page 371) wrote that little research on the effectiveness of group diagnostic tests in arithmetic has been reported. At the risk of being considered pretentious, it can be said that the results obtained in the development of this test demonstrate that it is effective in identifying systematic errors in subtraction and that it can be used successfully by classroom teachers. This effectiveness is due to the method by which the test was constructed and developed. Diagnostic instruments are typically constructed by first defining a hierarchy within a content domain, writing items corresponding to elements in the hierarchy and then administering the test. The process of diagnosis begins when the results are analyzed for error patterns and faulty concepts hypothesized. Verification is then attempted by means of interviews. The search for error patterns may be omitted and only the interview conducted. Faulty concepts are then hypothesized and tested during the interview.

Construction of this test started by a systematic review of the research on errors made in subtraction and logical analysis of the subtraction algorithm. These identified possible faulty concepts or misconceptions. Items were then devised to provide opportunity for these misconceptions to be applied and a list of error patterns which would result from their systematic application was generated. After administration of the test, interpretation consists simply of matching pupils' incorrect answers to the list of error patterns. The matching identifies the faulty concept. This is somewhat along the lines proposed by Denmark (1974), but he has published no further work in this area. The technical section describes the test development procedures and results.

Test Development

An extensive review of the literature relating to errors made in subtraction by elementary school pupils resulted in a list of over thirty types of errors which appeared to be worthy of further study. Analysis of those errors showed that a maximum of twelve types of computations were sufficient to produce these errors. These are listed below in algebraic form:

A. $\begin{array}{r} w \\ -\underline{a} \end{array}$	B. $\begin{array}{r} la \\ -\underline{w} \end{array}$	C. $\begin{array}{r} pw \\ -\underline{a} \end{array}$	D. $\begin{array}{r} pa \\ -\underline{w} \end{array}$
E. $\begin{array}{r} wx \\ -\underline{ab} \end{array}$	F. $\begin{array}{r} wa \\ -\underline{bx} \end{array}$	G. $\begin{array}{r} pwg \\ -\underline{ag} \end{array}$	H. $\begin{array}{r} pab \\ -\underline{wx} \end{array}$
I. $\begin{array}{r} wxa \\ -\underline{b0y} \end{array}$	J. $\begin{array}{r} w0a \\ -\underline{bpx} \end{array}$	$\begin{array}{r} w00a \\ -\underline{bpqx} \end{array}$	L. wxy-ab

In the exemplars above the digits represented by the letters a,b,c and d are less than the digits represented by w,x,y and z. Digits represented by the letters p,q,r and s may be any size. 0 is zero.

Types A and B test only number fact knowledge and are embedded in the other item types. Type C is also embedded in other item types. Therefore these three types were not retained as separate entities for further study. Type L was also discarded since, in practice, children will generally first change the horizontal format to a vertical format before starting the computation. This left eight types of computations.

The exemplars listed above were used to generate computations and a series of tests was then prepared. Each test was composed of twenty items, ten of each of two item types. Each test was then administered to separate samples of fourth, fifth and sixth grade pupils. Each sample comprised approximately 100 pupils. The tests were then marked and wrong answers coded. The coding system was based on the assumption that each of the error types previously identified would produce unique wrong responses and thus identify the faulty concept utilized by the child. Pupil working was analysed and follow up interviews conducted to determine whether the faulty concept indicated by the coding system was in fact the concept used by the child. From this procedure it was apparent that items generated by the same algebraic exemplars did not in fact necessarily generate similar errors. Consequently, selection criteria for items were devised as follows:

1. Error Production. Because of the nature of a diagnostic test, test items should provide opportunities for children who have faulty subtraction concepts to make errors. Two difficulty indices were computed for each item, one which considered all those children taking the test, and one which considered only those who made three or more errors on the test. Items were considered good if they were done corectly by 90% of all children taking the test, (difficulty index 0.9), were done incorrectly by 50% of these children making 3 or more errors, (difficulty index 0.5), and the item variance was maximized where item variance = difficulty index X (1 - difficulty index).

2. Uniqueness of Error. Interpretation of this test is built upon the assumption that each wrong answer points to a particular faulty concept. Thus it was necessary to avoid combinations of digits which provided opportunity for different faulty concepts to produce the same wrong answer. It proved impossible to avoid duplications completely, but they were minimized and the error code guide explains how to interpret in such situations.

3. Type of Error Produced. Each item type was devised to produce a small number of specific wrong answers. Those items which produced too many errors which could not be anticipated or explained by the theoretical basis for construction of these items were eliminated. Decisions were made on the ratio of codable errors (i.e., one of the predicted wrong answers) to uncodable errors, (i.e., without apparent systematic basis). Items included in the test were those with ratios of 65-70%. In later piloting of

the test itself ratios of 60-65% were obtained. Since many errors made by children are actually combinations of errors, this was considered good.

4. Comparison of High/Low Pupil Performance. On a diagnostic test pupils who have no faulty concepts should score higher than those children who have. This should also be true for each item. Therefore the performance of those pupils making fewer than three errors was compared with the performance of those making three or more errors, using the difficulty indices described above. Items selected were those that discriminated best between these two groups.

Application of these criteria resulted in a 24-item test containing three items of each of the three item types. This preliminary test was administered to over 900 pupils in 6 schools. The tests were scored and coded. The test papers of 140 children who were identified by the test as making systematic errors were chosen for further analysis of their working. Follow up interviews were also conducted with each of these children to determine whether the systematic errors and associated faulty concepts were accurately identified by the test. This led to clarification of error descriptions and the inclusion of a number of new categories of error. Several error categories were omitted because of extremely low incidence. At this stage it was recognized that item types D and E contributed very little to the predictive value of the test. Moreover the errors identified by these item types were also identified by other item types on the test. Therefore the Type D items were eliminated and one additional item was included for item types G, H and J. Those items added were drawn from the item pool previously tested and met the established criteria. The three type E items were retained as easy items to prevent the possibility of the test being too intimidating. The final version of the test now consisted of 24 items comprised of either three or four items of each of seven item types. Thus the test can be viewed as a composite of seven sub-tests which are to some extent independent.

This final version of the test, with its expanded number of error categories, was then administered to 522 fifth and sixth grade pupils, scored and the errors coded. The test papers of those pupils who made three or more errors (n = 355) were then subjected to item analysis. The item analysis was carried out usng a LERTAP computer program. Table 1 shows the percentage of correct and incorrect responses for each item and the mean scores on the subtest of those making correct or incorrect responses to that item. These percentages are appropriate for a diagnostic test. The means indicate that those making a correct response on an item tended to do well in the subtest while those who made an incorrect response tended to do poorly on the subtest. This means the test did identify those making a systematic error.

Table 1. Percentage Correct/Incorrect Responses and Sub-test Means

Sub-test Type	Item	Percentage		Sub-test Mean	
		Correct	Incorrect	Correct	Incorrect
E	1	97.3	2.7	2.86	1.16
	2	98.4	1.6	2.94	1.23
	3	96.9	3.1	2.95	1.21
F	1	94.9	5.1	2.95	1.22
	2	95.8	4.2	2.93	1.27
	3	95.2	4.8	2.93	1.35
G	1	98.6	1.4	3.95	2.20
	2	96.9	3.1	3.97	2.45
	3	97.5	2.5	3.96	2.33
	4	99.2	0.8	3.94	1.67
H	1	91.3	8.7	3.77	1.52
	2	90.4	9.6	3.79	1.56
	3	87.9	12.1	3.79	1.95
	4	87.6	12.4	3.81	1.91
I	1	90.1	9.9	2.87	0.80
	2	89.0	11.0	2.89	0.90
	3	87.6	12.4	2.89	1.11
J	1	90.4	9.6	3.82	1.56
	2	90.1	9.9	3.84	1.46
	3	89.9	10.1	3.86	1.50
	4	90.1	9.1	3.84	1.43
K	1	88.5	11.5	2.89	1.10
	2	89.6	10.4	2.88	0.95

Estimates of internal consistency are shown below in Table 2.

Table 2. Hoyt Estimates of Reliability for Each Sub-Test.

SUB-TEST	NO. OF ITEMS	HOYT COEFFICIENT	SEM	ITEM TYPE
1.	3	0.25	0.20	pa - w
2.	3	0.59	0.25	wa -bx
3.	4	0.54	0.21	pwg - ag
4.	4	0.68	0.43	pab - wx
5.	3	0.72	0.32	wxa -b0y
6.	4	0.77	0.38	w0a -bpx
7.	3	0.70	0.33	w00a -bpgx

Where (i) a,b,c,d < w,x,y
(ii) pqrs may be any size single digit.
(iii) 0 represents zero.

Cronbach's alpha =0.68.

Because of the fact that these reliability estimates are inevitably deflated due to the dichotomy (correct ✓ incorrect) and the low number of items in each sub-test, these figures are regarded as very encouraging. The one sub-test which scored low comprised those easy items which were already known to be poor predictors and to add little to the power of the test. At this point, then, a reliable test had been developed which could identify systematic errors and the faulty concepts which produced them. The question became one of whether teachers could and would use the test.

Classroom Usability

One of the primary goals in developing this test was to make it usable in the classroom. In order to test the usability a study was carried out which involved 24 teachers and 700 fifth and sixth grade students. Each participant was provided with a classroom test set which included the following:

a) A one page diagnostic test sheet for each child.
b) A master answer sheet.
c) A coding matrix for coding wrong answers.
d) A description of the error categories.
e) Guidelines for administering and interpreting the tests.
f) A diagnostic record.

These materials were delivered personally to all schools involved in the project. During this visit and in a cover letter to each teacher it was made clear that the researchers would be available if clarification or help were needed. No such requests were received.

Each teacher involved in the project was requested to follow the procedures described below. These are more fully explained in the guidelines for administering the test. These steps were:

a) Administer the test at the earliest opportunity.
b) Score the test.
c) Code all wrong answers according to the coding matrix supplied.
d) Make diagnoses of problem cases using the error descriptions given.
e) Analyse pupil working to determine whether the working supported the initial diagnosis.
f) Interview problem cases, i.e., cases of systematic error to determine whether the initial diagnoses were accurate.
g) Complete the diagnostic record.

When all the teachers had completed the above procedures all tests were collected and a questionnaire distributed. This questionnaire solicited information regarding time commitment and teacher attitude toward using the test. These were completed anonymously. Steps c through g above were then repeated by a research assistant (RA) trained in the test procedures for all the classes in the project.

Of the 700 test papers collected, 148 were found to have three or more errors and were subjected to further analysis. The remaining test papers were considered to be problem free and were excluded from the subsequent analysis.

Data analysis was based on errors, rather than subjects, and each error was treated as an independent observation. A total of 853 observations were recorded. In addition to demographic data, the following information was recorded for each observation:

a) The test item for which the error was made.
b) The error type as indicated by the error code.

These were followed by a trichotomous numeric code to indicate:

c) Whether the analysis of working supported the test diagnosis.
d) Whether the teacher interview supported the test diagnosis.
e) Whether the RA analysis supported the test diagnosis.
f) Whether the RA analysis supported the teacher analysis.
g) Whether the teacher interview supported the teacher analysis of working.
h) Whether the RA analysis supported the teacher interview.

This information was keypunched and analyzed using an appropriate SPSS program which produced frequency counts and contingency tables. As a further measure of agreement on items f, g and h, a series of Kendall's Coefficients of Concordance were computed. The teacher questionnaires were summarized manually.

Results. The results of the above analysis are presented in Tables 3 - 11.

Table 3. Teacher Analysis Support of Test Diagnosis

	NO	YES	TOTAL
n	109	392	501
%	21.8	78.2	100

In 352 cases of error teachers did not analyse the working and these cases are not reflected in the table above. This is partly explained by the fact that when number fact errors are made there is generally no working to analyse.

Table 4. Interview Support of Test Diagnosis

	NO	YES	TOTAL
n	85	209	294
%	28.9	71.1	100

In 559 instances the teacher interview was not done. The reasons for this are not known, but it appeared that where the errors were of number facts rather than process teachers tended not to interview.

Table 5. RA Analysis Support of Test Diagnosis

	NO	YES	TOTAL
n	163	512	675
%	24.1	75.9	100

In the 178 missing cases there was no working to analyse.

Table 6. Agreement Between RA and Teacher Analysis of Working

		RA		
		NO	YES	TOTAL/%
	NO	93	16	109/21.8
Teacher	YES	27	365	392/78.2
	TOTAL/ %	120/24.0	381/76.0	501/100

The table shows that in 365 out of 501 cases there was agreement. A corrected Chi Square (1df.) and a Phi were calculated to be 283.744 ($p < 0.0001$) and 0.7582 respectively. This shows significant agreement between the two measures.

Table 7. Interview Support of Teacher Analysis

		Interview		
		NO	YES	TOTAL/%
	NO	63	3	66/22.4
Analysis	YES	22	206	228/77.6
	TOTAL/%	85/28.9	209/71.1	294/100

A corrected Chi Square (1df.) and a Phi were calculated to be 179.20 ($p<0.0001$) and 0.7897 respectively. Clearly, this shows a significant correlation between the two measures. Note that the table includes only instances where an interview was carried out.

Table 8. Interview Support of RA Analysis

		Interview		
		NO	YES	TOTAL/%
	NO	64	14	78/26.5
RA	YES	21	195	216/73.5
	TOTAL/%	85/28.9	201/71.1	294/100

The table shows that in 195 cases out of 294 the teacher interview and the RA analysis of the working were in agreement. A corrected Chi Square (ldf.) and a Phi were calculated to be 142.36 ($p<0.0001$) and 0.7044 respectively. Again this shows significant agreement between the two measures.

Tables 9 - 11 present data relative to Kendall's Coefficient of Concordance.

Table 9. Agreement Between RA Analysis and Teacher Analysis

Item	No. of cases	W	Chi Square	D.F.	Probability
1	67	0.191	12.800	1	0.000
2	17	0.320	5.444	1	0.020
3	11	0.636	7.000	1	0.008
4	60	0.171	10.286	1	0.001
5	54	0.132	7.118	1	0.008
6	44	0.162	7.143	1	0.008
7	27	0.074	2.000	1	0.157*
8	71	0.214	15.211	1	0.000
9	34	0.053	1.800	1	0.180*
10	57	0.112	6.400	1	0.011
11	-	-	-	-	-
12	22	0.091	2.000	1	0.157*
13	27	0.185	5.000	1	0.025
14	67	0.198	13.235	1	0.000
15	59	0.203	12.000	1	0.001
16	31	0.145	4.500	1	0.034
17	15	0.067	1.000	1	0.317*
18	37	0.189	7.000	1	0.008
19	57	0.298	17.000	1	0.000
20	17	0.235	4.000	1	0.046
21	16	0.563	9.000	1	0.003
22	5	0.200	1.000	1	0.317*
23	42	0.268	11.267	1	0.001
24	12	0.0	0.0	1	1.000*

*non-significant ($p>0.05$)

Table 10. Agreement Between Teacher Analysis and Teacher Interview

Item	N	W	Chi Square	DF	Probability
1	67	0.205	13.762	1	0.000
2	17	0.235	4.000	1	0.046
3	11	0.273	3.000	1	0.083 *
4	60	0.150	9.000	1	0.003
5	54	0.296	16.000	1	0.000
6	44	0.234	10.286	1	0.001
7	27	0.296	8.000	1	0.005
8	71	0.088	6.250	1	0.012
9	34	0.235	8.000	1	0.005
10	57	0.112	6.368	1	0.012
11	-	-	-	-	-
12	22	0.364	8.000	1	0.005
13	27	0.259	7.000	1	0.008
14	67	0.198	13.235	1	0.000
15	59	0.092	5.400	1	0.020
16	31	0.145	4.500	1	0.034
17	15	0.120	1.800	1	0.180 *
18	37	0.097	3.600	1	0.058 *
19	57	0.180	10.286	1	0.001
20	17	0.235	4.000	1	0.046
21	16	0.250	4.000	1	0.046
22	5	0.400	2.000	1	0.157 *
23	42	0.130	5.444	1	0.020
24	12	0.083	1.000	1	0.317 *

* $p>0.05$

Table 11. Agreement Between RA Analysis and Teacher Interview

Iten	N	W	Chi Square	DF	Probability
1	67	0.425	28.444	1	0.000
2	17	0.706	12.000	1	0.001
3	11	0.909	10.000	1	0.002
4	60	0.369	22.154	1	0.000
5	54	0.484	26.133	1	0.000
6	44	0.401	17.640	1	0.000
7	27	0.333	9.000	1	0.003
8	71	0.345	24.500	1	0.000
9	34	0.353	12.000	1	0.001
10	57	0.253	14.440	1	0.000
11	-	-	-	-	-
12	22	0.455	10.000	1	0.002
13	27	0.444	12.000	1	0.001
14	67	0.420	28.125	1	0.000
15	59	0.299	17.640	1	0.000
16	31	0.363	11.267		0.001
17	15	0.178	2.667	1	0.102 *
18	37	0.269	9.941	1	0.002
19	57	0.458	26.133	1	0.000
20	17	0.471	8.000	1	0.005
21	16	0.813	13.000	1	0.000
22	5	0.600	3.000	1	0.083 *
23	42	0.374	15.696	1	0.000
24	12	0.150	1.800	1	0.180 *

* $p>0.05$

Questionnaire Results: A total of 20 teachers responded to the questionnaire and the results are summarized below. Not all teachers completed all items.

Table 12. Estimate of Time Taken

Activity	Time (Hrs.)		
	1	1-2	3
Total, excluding interview	8	5	3
Interviews	12	4	0
If repeated, excluding interviews	9	5	2
If repeated, interview time	12	4	0

Table 13. Teacher Opinion of Usability

Low ⟷ High

	1	2	3	4	5
Ease of Administration			1	9	9
Clarity of guide and error categories		2	2	10	15
Willingness to use			2	9	9
Value of test	2	0	4	6	6

These results are interpreted as being very positive. While the weakest category would appear to be that relating to the value of the tests, teacher comments on the questionnaire seem to indicate that the perceived value was related to the class ability. Teachers who uncovered a number of problems felt it was of more value than those whose classes made few errors.

Concerning classroom usability the following conclusions were drawn:

1. The initial diagnoses of errors as indicated by the test were strongly supported by teacher analyses of working, teacher interviews and by an independent observer (RA), This indicates that the test is effective in identifying systematic errors in subtraction, and that in most cases a diagnostic interview is unnecessary.

2. The high degree of agreement between teacher and RA analyses and teacher interview and RA analysis indicates that the test is a reliable measure when used by elementary teachers.

3. The teacher guide was generally seen to be clear and the test easy to administer.

4. The test was efficient in terms of time commitment and obtained results and teachers would be willing to use the test.

In short, the test appears to do what it was designed to do, and elementary teachers can and would use it.

References

Ashlock, Robert. Error Patterns in Computation: A Semi-Programmed Approach; 2nd Ed. Columbus: Charles Merrill Co., 1976.

Buros, O.K. (Ed.). Mathematics Tests and Reviews, Highland Park, N.J. The Gryphon Press, 1975.

Buswell, G.T., and Judd, C.H. Summaries of Educational Investigations Relating to Arithmetic. Chicago: University of Chicago Press, 1925.

Denmark, T. A reaction paper to Classroom Diagnosis by Robert Underhill. Paper presented at First National Conference on Remedial Mathematics, Kent, Ohio, 1974.

Hopkins, Martha H. Teaching Arithmetic: A Diagnostic Approach. Bradenton, Fl. Martha H Hopkins, 1977.

Reisman, Frederika. A Guide to the Diagnostic Teaching of Arithmetic. Columbus: Charles E. Merrill Co., 1972.

Speer, William R. A Clinical Model for Diagnosing Mathematical Deficiencies Incorporating Educational Cognitive Style. Paper presented at the Fourth National Conference on Diagnostic and Prescriptive Mathematics. University of Maryland, 1977.

Uprichard, A. Edward, et al. A Task Process Model for Diagnosis and Prescription in Mathematics. Paper prescribed at the Fourth National Conference on Diagnostic and Prescriptive Mathematics, 1977.

Note: Because of space constraints it was not possible to include the extensive appendix accompanying the paper.

An Exploratory Study of Mathematical Abilities: The Learning and Teaching of Some Basic Mathematical Abilities

Ruth Rees
Brunel University

During the last few years there has been concern, both nationally and internationally, over the performance of school pupils, college students and others in basic mathematical skills. The series of studies described in this paper were pioneering investigations which attempted, in a modest way, to quantify the nature of the problem at various educational levels. They were carried out in the Department of Education at Brunel University during the last decade under the overall supervison of Professor W.D. Furneaux. They originated in 1970 with a specific concern in the Further Education sector, but subsequently spread to secondary schools, teachers in training and university engineering undergraduates. Work of the Brunel Education Group has now been extended to primary schools and to more complex mathematical concepts and skills.

Essentially the following questions have been investigated:

1. Do school pupils and college students experience particular difficulties in the learning of mathematics?

2. If so, what is the nature and extent of these difficulties?

3. What are the reasons for these difficulties?

4. What can be done to resolve the difficulties?

In attempting to answer these questions the analyses of the data led to investigating the determinants of successful mathematical performance and thus the structure of mathematical abilities - in other words, towards the beginning of a theory for mathematics learning and teaching.

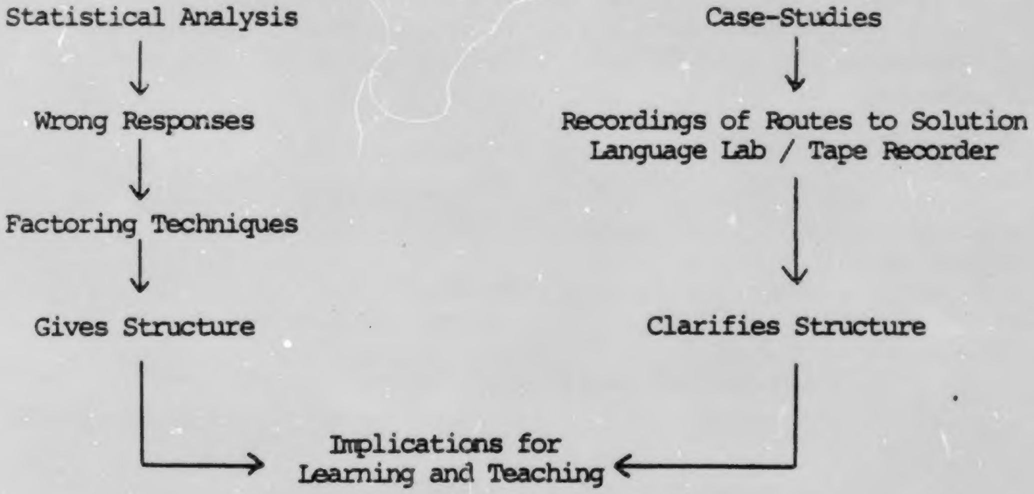
It was at this point that I had the pleasure of relating the analysis of considerable experimental data to Richard Skemp's work and currently to his new model of intelligence. As he points out in 'Intelligence, Learning, and Action', there is a need for a satisfactory theory on which to base the activities of learners and teachers i.e., a need for a unified body of knowledge. "Every new teacher has to learn mostly from his own mistakes" (p.248). What an inappropriate state of affairs for teaching and therefore learning! The late Ed. Begle in "Critical Variables in Mathematics Education", also laments the lack of a theory despite the volume of studies carried out.

This paper will be concerned with some of the main aspects of the Brunel studies. In England, compulsory education is from 5-16yrs. Depending upon academic achievement, students may continue in school then to a University or Polytechnic (Higher Education) or pursue various types of vocational courses (Further Education).

METHODOLOGY

An attack on the problems of learning and teaching mathematics needed a combination of approaches. The traditional paper and pencil method enables analyses to be carried out on a considerable quantity of test data. The more currently popular method of case-studies, on the other hand, gives insight into the perception and thought processes of the individual student. There has been unfortunate and unnecessary controversy over the relative merits of these two approaches. It was decided for the purposes of these studies to use these approaches in a complementary manner as shown in Figure 1.

Fig. 1. Complementary Research Methods .


Statistical Analysis
Case-Studies
Wrong Responses
Recordings of Routes to Solution
Language Lab / Tape Recorder
Factoring Techniques
Gives Structure
Clarifies Structure
Implications for
Learning and Teaching

DESIGN AND ANALYSES OF THE MATHEMATICAL TASKS

1. Design.

The construction of the set of mathematical tasks was based on the kinds of knowledge with which school leavers should be equipped. These consist essentially of arithmetical concepts and skills together with some of those of geometry and algebra. The acquisition of skills on these tasks would prepare the student for continuing education, employment or general ways of living.

The set was originally designed for further education craft students and was based on the analyses of the performance of a national sample of 17 000 craft engineering students on a new City and Guilds of London Institute examination. The curriculum for this examination was the result of careful consultations between representatives from industry, education and CGLI staff.

A total of 50 mathematical items was designed and for the purposes of the investigations was called a 'Calculations Test'. It was designed to be a power rather than a speed test so that, as far as possible, all students should attempt all items. Since it was to be a diagnostic test the multiple-choice format was selected in order to define the regions of difficulty and by suitably selected distractors (or attractors) it was planned to pinpoint the nature of the difficulties.

Although some of the items were in a craft context they were constructed to be independent of a knowledge of craft formulae: for example, in the items involving electrical craft, Ohm's law is given. Formulae for calculating the circumference and area of a circle, however, were not given , in order to test for difficulty with this topic.

Some of the items in the test are given in the appendix. It will be noted that items 8, 9 and 10 involve the same calculation in different contexts: pure calculation, algebraic, vocational.

2. Samples.

The set of tasks was originally designed for F.E. Craft level students but was subsequently given to students on F.E. Technician and O.N.C. courses and to a sample of university engineering undergraduates. As these students had recently left school, the tasks were also given to many samples of secondary school pupils from ages 14-17 covering a wide range of school organisation and mathematics syllabuses. The studies were also extended to teachers in training, primary and secondary and inservice practising teachers.

3. Analysis of wrong responses.

Item analyses were made for many kinds of subgroups of students according to age, type of mathematics course and so on. An item was defined to be difficult if 50% or more of a group of students could not do it correctly. For all subgroups it was found that there was a subset of items which gave persistent difficulty, the extent of which depended on the educational level. This subset was called the 'core of difficulty'.

Teachers in training also shared this core of difficulty. These teacher trainees were at the end of their second year on a 3 year certificate course and had just completed an end of course mathematics examination. They could therefore be said to be at the peak of their mathematical 'power'. The attitude of these trainees also gives cause for concern. This seems to indicate a learning-teaching cycle and the intention of the phrasing of the second half of the title may now be clearer.

4. Characteristics of the care of difficulty.

The following observations were made:

The concepts and skills tested by core items were found to include:

(i) Operations on numbers less than unity.
(ii) The Reciprocal/Inverse.
(iii) Mensuration of the circle.
(iv) Property of similar figures: Area α (dimension)^2 .
(v) Changing the form of numbers (e.g., fraction to decimal).

One example of each of these is as follows:

(i) The square root of 0.9 is:
(ii) If $\frac{1}{R} = \frac{1}{2} + \frac{1}{6}$ then R is:
(iii) The diameter of the base of a cone is 30mm . Area of base in mm^2 is:
(iv) If the diameters of two circles are in the ration 1:2 the ratio of their areas is:
(v) $\frac{7}{16}$ correct to 2 decimal places is:

These items form a heterogeneous set: there is no mathematical process common to all. They were judged by mathematicians and mathematics teachers to be apparently no different from the other items in the test with regard to construction. And yet, on these items:

(i) Performance is relatively worse for all kinds of students.
(ii) Students appear to need more mathematical "inference" for solution.
(iii) Teachers often do not appreciate the nature of the difficulties and underestimate the difficulties.
(iv) A learning-teaching cycle appears to be generated by these problems.

Questions were posed as to the nature of these items: could they, for example, represent some special kind of ability? In order to answer these questions further studies were undertaken and it was decided to use the techniques of factor analysis on the resulting data.

THE FACTOR ANALYTIC STUDIES

The factor analytic technique was applied first to the initial set of fifty items. The departure in this analysis from all previous studies was that each mathematics item was entered into the analysis as a variable. For all subgroups it was found that there were two main 'clusters' of mathematics items one of which was the core of difficulty. The other main cluster consisted of items which were judged to be more amenable to 'mechanical' solutions and for which teachers tended to make more accurate judgements of students' performance.

In order to find correlates with these clusters subsequent studies, using established psychological variables, were carried out. These included two models of general intelligence. The first was the Thurstone Primary Mental Abilities Battery consisting of Verbal, Numerical, Total Reasoning and Spatial Abilities. The second was the Nufferno model consisting of the three components of speed, accuracy and persistence (continuance). The results of these studies have been fully reported elsewhere. (7, 8).

It will be noted in table 1, that the 'core' items are relatively independent of the general ability 'g' measures whilst the other cluster correlates with the measures used for 'g'.

No definite correlates have yet been found for the 'core': the nearest are certain of the Nufferno levels of continuance. Subjective judgements have been made by mathematics students of similarity to some of the 'Krutetskii' problems which he claims differentiate the mathematically able pupils.

Thus so far the various kinds of analyses used (i) indicated ways in which different kinds of students appeared to be thinking of the various mathematical tasks and (ii) enabled a general structure to emerge upon which further hypotheses could be formulated.

The paper and pencil method, however, gives the end result of a student's thinking processes. It was felt necessary to monitor the original perception of the task by the student and the subsequent route to the solution obtained. The language laboratory and the tape recorder were used as was appropriate: the former has the advantage of investigating several students simultaneously whilst the latter was also used in a face-to-face dialogue situation when required.

See Table 1 on page 110

THE LANGUAGE LABORATORY STUDIES

These studies have illuminated several aspects of task solution not revealed by the paper and pencil situation. One of these is the way in which a student sees a problem and this can be quite different from that of the teacher or the way in which the teacher assumes the student is thinking. Another aspect is the way in which this original perception appears to highlight the whole route to solution. Yet a third and important aspect is the revealing of where the discontinuity in the thought process occurs i.e., the 'blockage'. Diagnosis of the 'blockages' is subtle: within the same topic the nature of the discontinuity is found only by posing certain kinds of questions. It is clear that when students are asked to solve mathematical problems by working 'out loud', the kinds of difficulty they complain of when attempting items demanding 'mathematical inference' are recognisably different from those associated with problems calling more on general intellectual ability. The essential difference between a 'g' type mathematical problem and an inference-type problem is that the appropriate mathematical operation has to be recognised for the former and deduced for the latter.

DEVELOPMENT OF THE TWO-FACTOR STRUCTURE

Studies are now being carried out at primary and early secondary schools levels and F.E. Technician levels to investigate the possible emergence and development of the structure and the associated "development of the blockages".

There is evidence of (i) the emergence of a two-factor structure at 10+ where the factors tend not to be so differentiated. There is clearer differentiation at later secondary levels as may be expected with specialisation, (ii) the 'blockages/errors' emerging at primary level and developing through secondary to F.E. Levels. This is shown in Table 2.

TABLE 2: VALUES OF RATIOS OF VARIANCE ACCOUNTED FOR BY THE 'g' FACTOR AND 'INFERENCE' FACTOR

Type of Student	Ratio g:I
Further Education Students (17-19 years)	2:3
Secondary Pupils (14-15 years)	2:1
Junior School Pupils (10-11 years)	5:2

Other factors, such as the role of affective variables and sex-related effects are also being investigated.

SUMMARY: DISCUSSION OF IMPLICATIONS FOR TEACHING

The characteristics of the two clusters of test items based on the factor analysis may be summarized as shown in Table 3.

TABLE 3: CHARACTERISTICS OF TWO CLUSTERS OF TEST ITEMS

M_g	M_I
1. Students perform reasonably well provided they recognise the problem.	1. Performance is relatively worse for all kinds of students. Solutions have to be deduced.
2. Teachers' estimates of students' performance are reasonably accurate.	2. (i) Teachers often do not appreciate the nature of the difficulties and underestimate these difficulties. (ii) A learning-teaching cycle appears to be generated by these problems.

cont'd

3. Students appear to solve algorithmically.	3. Students appear to need more mathematical 'inference' for solution.

M_g are items calling more on general ability

M_I are items which require mathematical inference

How does this structure relate to the Skemp model? As might be expected, it is clear from the recordings that both sets of items are attempted to be solved in a 'goal-oriented' way, that is, students seek a solution. The processes involved, however, appear to be different, and they are summarized in Table 4.

TABLE 4: CHARACTERISTICS OF THE PROCESSES USED IN THE SOLUTION OF TEST ITEMS.

INTELLIGENT ACTIVITY MATHEMATICS	
Speed, Accuracy, Characteristics; amenable to Routine Rote/Algorithmic Applications i.e. INSTRUMENTAL LEARNING	Reflective/Inferential/ Adaptable/Relational/ Intuitive Characteristics. Needs relational understanding before successful application/ skill. i.e. RELATIONAL LEARNING

The M_I mathematics would appear to correspond to Skemp's idea of a 'cognitive map'. There is a variety of routes to the goal state, and it is the job of the learner to construct cognitive maps which enable a route to the goal state to be selected. The M_g mathematics would appear to correspond to the 'single path' solution where the learner gets lost if he deviates.

Many questions may be posed. For example, do the two factors reflect the ways in which mathematics is taught? There is some evidence from recent work with teachers that M_I is not explicitly taught. As one teacher said while talking about numbers less than unity: 'I have to admit, that I had not really considered this as a separate problem from 'mixed decimals', that is, it had not dawned on me that 1.6 x 1.6 has less attendant difficulties than 0.6 x 0.6. Occasionally I have been called to a desk and been asked: "Where have I gone wrong?", when a correct answer has been obtained. Further questioning has produced: "I have multiplied, but I have got a smaller answer than the number I started with." This has given rise to a short class-discussion on what happens when we multiply by numbers less than one, but I have never set out to teach this explicitly in the context of decimal fractions.

Is the structure to do with the nature of mathematics? The symbolism and structures of the subject make it specifically and almost wholly inferential. We therefore demand more inferential thinking at a much earlier age than we do for any other subject except say physics which we start to teach later anyway. Most pupils can get by in most subjects with a 'g' type thinking i.e., straight application of straightforward facts. We do not appear to demand inferential thinking in other subjects such as history until much later in the learner's life.

Are the M_I processes unique to mathematics or are they common to many disciplines at later stages of study?

The implications of these kinds of questions are fundamental, embracing all we seek to do in education. We should aim to:

(i) teach for understanding i.e. searching for relationships, generalisations, etc. to help bring about the concept.

(ii) practice the understanding in all its varied aspects - realistic/concrete/numerical modes. Skemp (p.258) points out that if the examples are "well chosen, with plenty of variability and increasing noise, there should be no chance of this kind of practice degenerating into rote memorizing".

(iii) 'routinize' the processes i.e. They appear to have M characteristics but there is a basic understanding "so that conscious attention is freed for the novel and problematic" i.e. for the next layer of conceptual complexity.

In short, we should teach so that $M_I \longrightarrow M_g$

REFERENCES

1. "Mathematics in Further Education. Difficulties experienced by Craft and Technician Students", Ruth Rees, Hutchinson Educational Ltd. February 1973.

2. "An Investigation of Some Common Mathematical Difficulties experienced by Students", 'Mathematics in School' (Mathematics Association), Ruth Rees, Vol. 3, January, 1974.

3. "Mathematics in Teacher Training Institutions. Some Difficulties Experienced by Teachers in Training", Ruth

Rees, I.M.A. Bulletin, December 1976.

4. "Intelligence, Learning and Action", Richard R. Skemp, 1979, Wiley & Sons.

5. "Critical Variables in Mathematics Education". Findings from a Survey of the Empirical Literature, National Council of Teachers of Mathematics, E. Begle, 1979.

6. "The Factorial Nature of Ability in Elementary Mathematics", J. Wrigley, Brit. J. Educ. Psychol., 28, 61-78, 1958.

7. "The Dimensions of Mathematical Difficulties". A set of problems proving more difficult than teachers expect at all levels in the educational system. W.D. Furneaux and Ruth Rees, Brunel University Department of Education. Occasional Publications Series No. 1, September 1976.

8. "The Structure of Mathematical Ability". W.D. Furneaux and Ruth Rees. Brit. J. Psychology, November 1978.

9. "The Psychology of Mathematical Abilities in Schoolchildren", V.A. Krutetskii, The University of Chicago Press, 1976.

APPENDIX

A SELECTION OF TEST ITEMS

1. $\frac{7}{16}$ correct to 2 decimal places is:

   a. 0.43   b. 0.44   c. 0.49   d. None of these

2. $\frac{9}{16} + \frac{5}{64}$ is:

   a. $\frac{13}{16}$ b. $\frac{23}{32}$ c. $\frac{43}{64}$ d. $\frac{41}{64}$

3. The square root of 600 is between:

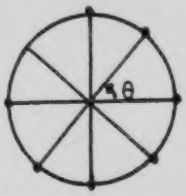
   a. 10 and 20   b. 20 and 30   c. 30 and 40   d. 40 and 50

4. The square root of 0.9 is:

   a. 0.3   b. 0.81   c. Between 0.9 and 1.0   d. None of these

5. If the diameters of two circles are in the ratio 1 : 2 the ratio of their areas is:

   a. 1 : π   b. 1 : 2   c. 1 : 4   d. 2 : 3



6. In the figure above, if the points marked on the circle are equally spaced then angle θ is:

   a. 30° b. 45° c. 60° d. None of these

7. For an increase in temperature of 1°C a 1 mm length of steel will expand 0.000011 mm. For an increase in temperature of 5°C, a 20 mm length of steel will expand by:

   a. 0.000055 mm   b. 0.00022 mm   c. 0.0011 mm   d. None of these

8. $8 \div 0.16$ is:

   a. 0.5   b. 2   c. 20   d. 50

9. If $0.16x = 8$, then x is:

   a. 1/2   b. 2   c. 5   d. None of these

10. Given that Resistance $= \frac{\text{Potential difference}}{\text{Current}}$, an appliance carrying a current of 0.16 amperes with a potential difference of 8 volts has a resistance of:

    a. 1.28 ohms   b. 1/2 ohm   c. 50 ohms   d. None of these

TABLE 1: SPSS FACTOR LOADING MATRIX FOR ROTATED FIRST SIX FACTORS (n = 255)

Item	F%	'core' I'(22.5%)	'g' II'(16.6%)	III'(4.3%)	IV'(3.4%)	V'(3.9%)	VI'(4.5%)
3	91.0	0.73					
4	72.9	0.58					
7	34.5	0.70					
11	51.4	(0.27)		0.44			
17	42.7			0.67			
18	17.6	0.49			0.60		
19	52.5	0.70					
20	20.4	0.61					
21	22.0			0.56			
22	32.2	0.63					
24	63.1	0.56					
26	50.2	0.61					
30	87.1	0.81					
31	24.3	0.78					
32	22.0	0.34		0.32			
33	65.1	0.63					
34	13.7	0.63					
37	44.7	0.63					
38	25.1	0.76					
44	11.8				0.73		
45	15.7			(0.29)			
10	82.0		0.43				
13	57.3		0.51				
28	34.1		0.38				
29	63.9		0.60				
39	80.4		0.40				
41	52.2		0.36				
48	58.0		0.36				
50	71.4		0.32				
51	71.0		0.52				
54	33.3						0.57
59	40.8					0.71	
60	40.0						0.54
62	72.2		0.44				
64	38.8					0.59	
66	50.6		0.35				
68	46.3		0.37				
69	43.9		0.46				
V			0.63				
N			0.73				
R			0.78				
S			0.62				
8	72.9	0.46	0.33				
25	25.9	0.35	0.30				
27	74.1	0.54	0.42				

The percentage variance accounted for by each factor is shown in brackets.
Facility value 'F' is the percentage of students selecting the correct solution.
Loadings < 0.30 are omitted except for items 11 and 45.

Mistakes in Mathematics - The Research of CSMS

Kathleen Hart
Chelsea College, University of London

The research project 'Concepts in Secondary Mathematics and Science' was financed by the Social Science Research Council for the five years 1974-79. The aim of the mathematics team was to investigate secondary school children's understanding of mathematics and to draw up a hierarchy of understanding.

To do this, the team constructed a test based on eleven topics which currently appear in the British secondary school mathematics curriculum. The topics were Algebra, Ratio, Measurement, Graphs, Fractions, Number Operations, Place Value and Decimals, Reflections and Rotations, Positive and Negative Numbers, and Vectors and Matrices. For each topic word problems were written. Very few computational items were included and technical terms and 'jargon' were kept at a minimum. Usually terms such as 'perimeter' was defined and conventions such as $x \rightarrow 3x$ were illustrated. These problems were field-tested by interviewing over 300 children in each topic area. The children were asked to do the questions and describe how they were seeking a solution.

After these trials the problems were put in the form of written tests which were administered to 10,000 children aged from 12+ to 15+ selected from 50 volunteer schools in England. The sample from each age group on each test was chosen, on the basis of non-verbal I.Q. tests, to represent the normal distribution of I.Q. scores in the British child population. The test papers were then marked and hierarchies in each topic drawn up based on item difficulty, inter-item correlation and degree of dependence between items.

The results of this study have been fully described elsewhere, (Hart, 1980) , but the interviews designed to elicit problem solving methods revealed that many errors were due to an identifiable incorrect strategy. Indeed, in the large scale study a number of errors were made by as much as fifty percent of the sample tested. Some of these are described below.

The methods used by children during the interviews showed why certain items were successfully completed by large numbers of children in the survey. Certain "child methods" which do not appear to have been taught formally in class but nevertheless appeared in a number of schools were often in evidence. These methods were often correct but limited. Consider the example of the following Ratio question used on interview with 17 children.

Onion Soup Recipe for 8 persons
8 onions
2 pints water
4 chicken soup cubes
2 dessertspoons butter
1/2 pint cream

a) I am cooking onion soup for 4 people.
   1) How much water do I need? ....
   2) How many chicken soup cubes do I need? ....

b) I am cooking onion soup for 6 people?
   1) How much water do I need? ....
   2) How many chicken soup cubes do I need? ....
   3) How much cream do I need? ....

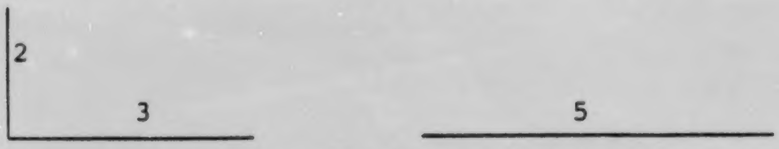
The most popular method of solving parts a(1), (2) was to take a half of the ingredients for eight people - not find the amount for one person and then multiply. To find the corresponding requirements for six people most of the children said 'half the amount for four added to the amount for four". This made part b(3) considerably more difficult than might have been expected if the child had already been multiplying by 3/4 because the "building up" method now resulted in 1/2(1/2) + 1/2(1/2 of 1/2) and as one child said "Is there another name for both of it?" About 20 percent of each year group tested on the written paper containing this question did not cope with the fraction but gave the answer 1/3 which is logical if one is trying to find a fraction greater than 1/4 but less than 1/2.

The fact that children regard ratio as an essentially additive process has also been stated by Lunzer (1966). The process is adequate for a number of questions; enlarging an amount in the ratio 5:2, for example, becomes "take it once, take it again, take a half and do an add". The method is not so easily applicable however when the amount is to be made smaller, for example:-

> An eel (25 cm in length) is fed a fishfinger 10 cm long; how much should another eel, 15 cm in length be fed if their appetites depend on their length?

Although this was the seventh question in a series, there was considerable hesitation over the answer. When asked "Why are these more difficult?" Brian replied "They're the minus ones. That's bigger so you really have to do minus to get there but you could share". The series of questions were adapted from tasks used by Piaget (1968) and of the 27 children interviewed on the eight parts (there were variations in the length of eel or which one's food was stated) only two children used the same method throughout. Even those who ultimately multiplied by a fraction, used an additive method when it was at all feasible.

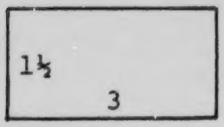
On the Ratio paper the hardest question was the enlargement of a figure in the ratio 5:3; it was on this question and three similar items (all of which dealt with enlarging figures although the other three required the ratio 3:2) that children used an incorrect addition strategy (identified by Karplus 1975). The hardest question was:-



Draw the new upright so that the right hand figure is the same shape but larger than the left.

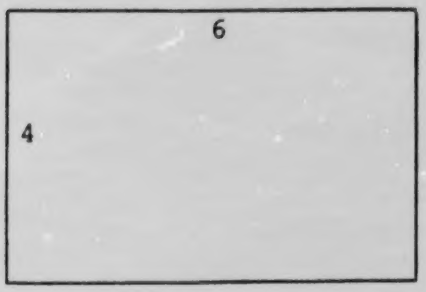
The addition strategy argument is "the base lines differ by 2 cm, so the new upright must be 2 cm longer than the original, answer 4 cm". This is not a building up method using addition as in the recipe question but it appeared sensible to about 40 percent of the sample tested on Ratio (n = 2257, ages 13, 14, 15). It may be significant that it appeared in the questions which involved the enlargement of figures and not in those which required the sharing of an amount "fairly". Unless the child draws the figure which results from the 'addition strategy' and it is grossly distorted it is very likely that having made the sides longer is considered a perfectly adequate enlargement. Take for example this interview with a girl of 14, asked to enlarge a rectangle so that it is the same shape but larger than the original.

1)



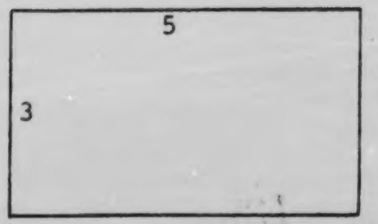
Enlarge (draw the same shape but larger)

2)



Enlarge with new base line 9 cm.

3)



Enlarge with new base line 12cm.

(Child: A Interviewer: I)

Question 1 - doubled.

A - I doubled that, I wasn't given a new line. If it was double it would have been 12. So its only 9, they added 3 on. In doubling its adding the same number again. I would have done that there if you hadn't given me that line! You can't have six times something to give 9. You can plus it. Goes on to Qu.3. Does by adding.

A - No the method is wrong. That's more of a square.

I - Reintroduces question 2. Can you times 9 by anything to give 3?

A - No

I - Is there a fraction?

A - much later - 1/3

I - Now 6 to 9?

A - Its going to be big, 1/2 or 3/4

I - writes 6 x - =9. Is it going to be more than 1?

A - Can I have whole numbers as well?

Writes 6 x 2 1/3 = 9

The statement 'You can't have six times something to give nine', is interesting because similar replies occurred in interviews on a number of mathematical topics. Children in England are traditionally introduced to fractions in the primary school at about the age of nine and by the age of fifteen are expected to work with the four operations and be at ease with fractions and decimals. This is far from the case however. It seems that possibly half our secondary school population is even at the age of fifteen firmly fixed within the set of whole numbers and the rules which they remember apply to those numbers e.g. "You cannot divide a smaller number by a larger". The following examples illustrate this.

1. 16 ÷ 20

AGE	12+	13+	15+
Fraction or decimal answer	7		36
0.16	4		6
Impossible	51		23

2. 3 ÷ 5

AGE	12+	13+	15+
3/5 or .6	31	35	
5/3	8.7	3.3	
1 2/5	9.4	5.3	
1 Rem. 2	17.5	18.3	

The responses for each age group are expressed as a percentage in all the examples given.

The rules for fractions are not remembered and at least 15 percent of each of the age groups 12+ to 15+ tested by CSMS (n = 1000) thought the rule for adding fractions was to add numerators and add denominators. Each child in this sample was given two forms of a test paper on fractions, one which contained problems and the other a series of computations which were thought to mirror the problems. Often the problems seemed easier than the computations; in particular the error just mentioned occured less when the problem form was given e.g.

1) 2/8 + 3/8

(computation)

Age	12+	13+	14+	15+
Correct (%)	77.6	66.3	71.8	67.9
Answer 5/16 (%)	8.5	19.7	14.3	16.7

2) In a baker's shop 3/8 of the flour is used for bread and 2/8 of the flour is used for cakes. What fraction of the flour has been used?

Age	12+	13+	14+	15+
Correct (%)	77	73	80.8	77.2
Answer 5/16 (%)	4.9	7.4	6.8	6

It often appeared as if the child thought he was dealing with two different topics. Faced with computations he asked himself 'what is the method?' whereas with the problems he possibly tried to find a meaning for the question. Certainly although the children did the computation paper immediately after that containing problems, they sometimes saw little connection between the two. For example the following four questions were all done within half an hour by the same children.

Age	12+	13+	14+	15+
3 x 10 1/2 (computation)	61.4	61.8	79.9	84.2
40 ÷ 10 1/2 (computation, 3 17/21 or 3 rem 8 1/2 counted correct)	25.2	20	20.8	27.4
(80/21 or 3 17/21 only answer accepted)	8.1	9.4	8.8	16.7
How many bicycle spokes 10 1/2cm long can be cut from a piece of wire 40cm long?	69.5	74.0	77.6	79.1
What length of wire is left over?	52.8	56.3	64.6	65.2

It seems fairly obvious that questions 3 and 4 were not seen as problems requiring the operations of multiplication and division but of addition and subtraction. There was often a lack of willingness (on the part of many secondary school children interviewed) to use multiplication when addition would suffice even though cumbersome.

Observations such as these have obvious implications for teaching. It seems likely that children do not assimilate a new method (or algorithm) if they find that they can manage (even if only for a limited period) with their own essentially additive methods. It also seems likely that numbers other than the natural numbers have little meaning for over forty percent of our secondary school children. This is possible because the models they are trying to use for fractions and decimals are those derived from whole numbers. If a child sees 3 x 4 as 4 + 4 + 4 then 3/5 x 4/7 has little meaning for him. If a teacher 'corrects' an error in class it is usually in the form of a reiteration of a procedure based on what he sees is needed. It is likely that the child never saw the question as a multiplication for example, so correcting a multiplication error when the child used addition is of little use to him.

Based on the interviews and the large-scale study the CSMS team has formulated hypotheses on which methods lead to certain errors, and a new research project, Strategies and Errors in Secondary Mathematics has been funded for 1980-83. This project is designed to look in depth at these methods and errors and to formulate teaching modules to remedy or modify them.

References

CSMS Mathematics Team (Editor K. Hart) Children's Understandng of Mathematics (11-16); John Murray, London, 1980.

Hart. K.M. Secondary School Children's Understanding of Mathematics (Research Monograph); Mathematics Education, Centre for Science Education, Chelsea College, London University, London, England, 1980.

Karplus, R., Karplus, E., Formisano, M., Paulsen, A.C., Proportional Reasoning and Control of Variables in Seven Countries; Advancing Education Through Science Oriented Programs, Report ID-65, June 1975.

Lunzer, E.A., Pumfrey, P.D. Understanding Proportionality; MATHEMATICS TEACHING, Vol. 3, Spring 1966.

Piaget, J., Grize, J.B., Szeminska, A., Bang, V. Epistemologie et Psychologie de la Fonction; Presses Universitaires de France, Paris, 1968.

Difficulties and Errors in Computation: A Report and Commentary on Recent Studies in England, Canada, and the United States

Douglas H. Crawford
Queen's University

Assessment has played an increasing part in education since the creation in 1964 of the Exploratory Committee on Assessing the Progress of Education in the U.S.A. which resulted in the project known since 1969 as the National Assessment of Educational Progress. This trend is due to a concern for basic skills and the need to respond to public concern that school standards were deteriorating.

In this paper seven assessment projects will be described and test results relating to computation with whole numbers, fractions and decimals reported and discussed, together with the educational implication of these findings.

The seven projects are:

1. First National Assessment of Educational Progress (NAEP I)	1972-3	USA
2. Second National Assessment of Educational Progress (NAEP II)	1977-8	USA
3. British Columbia Assessment	1977	Canada
4. Assessment of Performance Unit (APU)	1978	England
5. Schools Council Primary Mathematics Evaluation Studies	1972-5	England
6. Brunel University Diagnostic Studies of Basic Difficulties	1970-8	England
7. Institute of Mathematics and its Applications Pilot Test of Basic Numeracy (IMA)	1977	England

The NAEP and APU are continuing and national in scope, while the British Columbia project rivals them in size. All were aimed primarily or entirely at school age populations with the exception of the Brunel study which focussed mainly at post secondary students. Table 1 summarizes the essential characteristics of these projects.

See Table One on page 122

Difficulties and errors in Compution

A. Whole numbers

The coverage of items of variable difficulty is much better for the North American assessments than for the English projects. At present data is available for the A.P.U. Assessment only from newspaper reports, and for the other three projects the tendency has been to ask questions at a specific level so that no profile of development emerges.

For this reason and since the main purpose of this paper is to focus on errors, reference to actual performance levels is omitted. Tables 2-5 list the types of computation required of students tested in each of the seven projects, together with information on the errors observed.

See Tables 2 - 5 on pages 124 - 127

B. Common Fractions

In most of the projects being considered, items involving common fractions appeared in problems or in relation to decimal fractions. Additionally, information on difficulties and errors was relatively sparse, most of it relating to addition and multiplication of fractions. The information obtained is summarised in Tables 6, 7 and 8. No questions were located dealing directly with division of fractions.

See Tables 6 - 8 on pages 128 - 130

C. Decimals

Computing with decimals involves the extension of the place value concept to numbers less than unity. Major errors in addition and subtraction of decimals relate to this. Data for multiplication errors are sparse. Alignment of decimal point is less important for the actual multiplication but misalignment of the decimal point is still a noticeable error. In the Brunel University Project, special attention has been given to multiplication involving two decimals, each less than unity, especially when the transition to a zero in the first decimal place occurs.

There are almost no data available for division, the main error reported in the N.A.E.P. II study being a misplaced decimal point.

Information obtained is summarized in Table 9.

Discussion of findings.

The patterns of difficulties and errors seem to be pervasive. Many of them are well known and described in such works as that of Ashlock (1) and Suydam (2). The causes of these errors are not clear. According to Ginsburg (3), children's errors result from organised strategies and rules. He argues that for many children arithmetic as an activity is isolated from their ordinary concerns, and that there is a large gap between their informal, and concrete, and their formal, and abstract, knowledge. The drop in performance when a dollars and cents addition is replaced by the identical addition of numbers alone illustrates well this claim.

More recently, Radatz (4) has argued the need for developing cognitive models of the causes of errors, and produced an information-processing classification of them by consideration of the mechanisms used in obtaining, processing, and reproducing the information contained in mathematical tasks. Among his categories are errors due to:

(a) language difficulties;
(b) difficulties in obtaining spatial information;
(c) deficient mastery of prerequisite skills, facts and concepts, and
(d) the application of irrelevant rules or strategies.

Almost all of the major errors and difficulties found in the present survey fit into one of these four categories. Most are knowledge or process errors. At least two have spatial components, and certainly language difficulties are another continuing source of errors and incorrect responses.

In the last analysis, it is the teacher who has the greatest influence on learning. In the studies by Ward (5) (Schools Council) and Rees (6) (Brunel), the views of teachers on errors were also obtained. Thus teachers in the Schools Council's study gave errors in subtraction (reversal, problems with zero, and regrouping) and notation and place value as by far the most prevalent. Again, these are a mixture of knowledge and process errors. In the Brunel study, teachers were asked to estimate the difficulty level of each question. One of the significant findings was that experienced teachers underestimated the nature and extent of the difficulty of items identified as difficult by the groups tested. Those items were in general of a higher order of sophistication than the basic computations forming the subject of this paper. Yet they involve for their correct solution basic concepts and computations. Further analysis of the hierarchies involved seems strongly indicated. Research along the lines of a new approach to assessment recently reported by Ekenstam and Nilsson (7) from Sweden seems very promising. The approach aims to reveal difficult steps in the learning processes of selected topics by starting with a sophisticated question and constructing related questions formed from it in sequence by removing each final step or detail from the previous question. For example, given the question $\frac{3x - 2}{2} = \frac{x}{3}$,

a sequence of problems can be constructed corresponding to each step in a particular solution.

$$\begin{aligned} 3(3x - 2) &= 2x \\ 9x - 6 &= 2x \\ 7x - 6 &= 0 \\ 7x &= 6 \end{aligned}$$

is such a sequence.

Conclusion

Assessment at the national or provincial level has both advantages and disadvantages. From the point of view of improving both teaching and learning, assessment needs to be viewed diagnostically. It is a source of a very large body of data which can be examined to see in which areas learning is effective or ineffective. But this is only a starting point. It suggests avenues of research, and areas where teachers must do more thinking about the formation of concepts, the retention of facts, and the accurate development of skills. In doing so, one crucial factor is the need for more 1 - 1 clinical and diagnostic interviews.

Errors are better prevented than remediated. On the one hand we must continue and refine our research; on the other hand, we must communicate the necessary positive attitude, skills, and knowledge to teachers, both pre-service and in-service, more effectively, so that they can halt the errors at their source.

Note: The extensive bibliography submitted by the author was, for considerations of space reduced to these sources specifically cited in the discussion.

References

1. Ashlock, Robert B. Error Patterns in Computation. A semi-programmed approach. Charles E. Merrill Publishing Company, Columbus, Ohio, 1976.

2. Suydam, M.N. Review of Recent Research Related to the Concepts of Fractions and of Ratio. Procedings of the Second International Conference for the Psychology of Mathematics Education, Osnabruck, 1978.

3. Ginsburg, Herbert. Children's Arithmetic: The Learning Process. D. VanNostrand Company, New York, 1977.

4. Radatz, H. Error Analysis in Mathematics Education Journal for Research in Mathematics Education, Vol. 10, No. 3, May, 1979, pp. 163-171.

5. Ward, Murray. Mathematics and the 10-year old Schools Council Working Paper 61, Evans/Methuen Educational Press, London, 1979.

6. Rees, Ruth M. Some Basic Mathematical Difficulties Experienced by Students and the Reasons for their Difficulties. Proc. Third Instructional Congress on Mathematical Education, 1977.

7. Ekenstram, A.A., and Nilsson, M. A New Approach to the Assessment of Children's Mathematical Competence. Educational Studies in Mathematics, 10, 1979 pp. 41-46.

TABLE 1: OBJECTIVES, SCOPE AND SCALE OF SELECTED ASSESSMENT STUDIES IN U.S.A., CANADA AND ENGLAND

PROJECT	MAJOR GOALS	DEVELOPMENT OF AND RATIONALE FOR MATHEMATICS OBJECTIVES	TESTING PROGRAM			EXERCISES
			BASIS AND SAMPLING		STUDENTS TESTED	
1. N.A.E.P. 1st ASSESSMENT U.S.A.	Measure educational attainment; identify problem areas and educational priorities.	A. Objectives developed by E.T.S. and Psychological Corp. of America. Reviewed by lay citizens. Final revision by P.C.A. in 1970. B. Categorised by 3-dimension scheme of 1) uses 2) content 3) abilities. C. Exercises developed using 17 content areas and 6 abilities (recall; manipulate; understand P/S; use; appreciate).	Criterion referenced; specific task presented and percentage able to answer it estimated.	Probability matrix sampling. About 1% of each year group.	9yr 13yr 17yr ADULT 25000 30000 33000 4500	Packages used; mainly group, but some individual; tapes and reading aloud used to minimise lack of reading ability. Most open-ended.
2. N.A.E.P. 2nd ASSESSMENT U.S.A.	As above. Monitor changes using baseline of 1st assessment.	A. Objectives developed by N.A.E.P. B. Categorised by 2-dimension scheme of 5 content areas and 4 cognitive processes (knowledge; skills; understanding; application). C. Additional topics assessed were attitudes, hand calculators, computer literacy and metrication.	Same as above	Same as above	17000 27000 27000 ----	Packages used; testing time 45 min., half M/C, half open-ended.
3. BRITISH COLUMBIA ASSESSMENT 1977	Measure and disseminate present state of learning in essential concepts and skills; use to improve learning.	A. Objectives developed by contract team and studied by management committee, then reviewed by panels comprising teachers, educators and lay people. B. Categorised by 3 domains (computation and knowledge; comprehension; and applications) split into more specific content objectives (between 10 and 12 according to grade level).	Criterion referenced	100% intended 87% realised	NUMBERS/GRADES TESTED 4 8 12 35277 42250 23166	Nearly all M/C. 60 minutes testing time; 60-72 items. Group administered.
4. ASSESSMENT OF PERFORMANCE UNIT, ENGLAND. (CREATED 1974) 1st ASSESSMENT 1978	Appraise present assessment instruments and methods; create new ones; identify and disseminate significant findings about achievement; monitor changes over time.	Objectives developed by monitoring team at N.F.E.R. advised by a steering committee. Categorised by 2-dimension models with 6 activities (concepts, skills, applications, generalisation and proof, investigations and creativity, and attitudes) and 5 content areas. N.B. Items on first three activities predominated in 1978 assessment.	Criterion referenced	2% per year group	STUDENTS TESTED 11 yrs 15 yrs 12000 (May 1978) 12000 (Nov 1978) May add 9 and 13 year olds later.	About 50 items per written test. Also practical and attitude tests on sub-sample (10%). Testing time 45-80 minutes

BEST COPY AVAILABLE

TABLE 1 (Cont'd)

PROJECT	MAJOR GOALS	DEVELOPMENT OF AND RATIONALE FOR MATHEMATICS OBJECTIVES	TESTING PROGRAM		EXERCISES
			BASIS AND SAMPLING	STUDENTS TESTED	
5. PRIMARY SCHOOL MATHEMATICS EVALUATION STUDIES SCHOOLS COUNCIL UNIVERSITY OF READING, ENGLAND. 1972-75	Review present state of primary mathematics. Investigate children's competence and teachers' attitudes and opinions and their relationship.	Framework of 7 content areas was developed by project team, and each broken down into a hierarchy of concepts. List of suitable topics in the mathematics curriculum of the 10 year old was then compiled and questions composed with the help of many teachers.	Not known; stratified random sample by size of county (six) and county borough (six), and size of school within each of 12 LEA's chosen. 40 out of 48 schools chosen participated.	2300 10 year olds.	Paper and pencil test. Four parallel versions of 27 questions.
6. DIAGNOSIS OF SOME BASIC MATHEMATICAL DIFFICULTIES EXPERIENCED BY STUDENTS AT VARIOUS EDUCATIONAL LEVELS BRUNEL UNIVERSITY, ENGLAND. 1970 ONWARDS	Investigate concern about mathematical performance of further education students Quantify and identify the nature of the problem across the educational spectrum. Investigate reasons for difficulties.	Analysis of over 17000 City and Guilds of London craft exam papers. Design of 50-item diagnostic M/C power test based on Part I craft studies exam, involving 7 areas of "content": numbers and basic operations (13) area (18) change of subject of formula (5) same calculations in different context (5) algebra (3) geometry (4) other items (7) (5 items overlapped)	Students from: a) 4th, 5th and 6th year in secondary schools. b) Craft I and Technician I further education courses. c) Teacher trainees from College of Ed., Tech. Ed. and Post-Grad. training depts. d) First year university. Relation to populations not clear. Total sample of 2207.	STUDENTS TESTED S.S. 759 F.E. 709 T.T. 601 U. 138	Written 4-option M/C. Calculations test of 50 minutes (or longer if needed).
7. PILOT TEST OF BASIC NUMERACY INSTITUTE OF MATHEMATICS AND ITS APPLICATIONS, ENGLAND. 1977	Provide quantitative evidence on the level of attainment in simple mathematical skills of pupils able to leave school in 1978.	Test developed by small committee of experienced mathematics teachers: a) 6 2-part questions on very simple mathematics, b) rest applying simple arithmetic skills to everyday situations.	6 L.E.A.'s representative with respect to location and population: each nominated 10 schools to be geographically representative.	STUDENTS TESTED 4th yr sec 1271 5th yr sec 7156	17 questions (8 2-part calculations; 9 problems). Test of 40 minutes maximum

BEST COPY AVAILABLE

TABLE:2 DIFFICULTIES AND ERRORS IN ADDING WHOLE NUMBERS

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I	4	From 26 to 24 type+43 type 31 48 +25 2-digit numbers only. (One horizontal format).	Failure to regroup Place value errors Format difficulty	8% maximum
2. N.A.E.P. II	5	From 21 to 4285 +54 3273 +5125 (One in word format).	Used subtraction Carrying errors Use of technical vocabulary	
3. BRITISH COLUMBIA	4 Gr. 4 including 1 Gr. 8	16 56+37 185 +7 +412 429 35 +531	Place value errors Used subtraction Failure to carry	up to 2% 2% 1%
4. A.P.U. 1977-78		Data not so far available		
5. SCHOOLS COUNCIL 1972-75	2	a) 349 +264 b) 238 +375	Errors in basic facts Errors in carrying	
6. BRUNEL UNIVERSITY	2	a) 8+3+5 b) 3 11 + 7	Data not available	
7. I.M.A. NUMERACY	2	a) 14+35 b) 43+282		

BEST COPY AVAILABLE

TABLE:3 **DIFFICULTIES AND ERRORS IN SUBTRACTING WHOLE NUMBERS**

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I 1972-73	4	From 36 to 1054 -19 - 865 All vertical format	Reversal error (9 yr. olds) One error in regrouping	12-22% 5%
2. N.A.E.P. II 1977-78	6	From 48 to 504 -17 -236 No regroupings to two regroupings includes one horizontal word format (hardest)	Reversal error (9 yr. olds) Regrouping errors	8-32%
3. BRITISH COLUMBIA	5 Gr. 4 including 1 Gr. 8	From 36 to 1054 -19 - 865	Reversal error Used addition wholly or partically	up to 8% 6%
4. A.P.U. 1977-78		Data not so far available		
5. SCHOOLS COUNCIL 1972-75	2	a) 348 b) 439 -193 -284	Reversal or carrying error	11%
6. BRUNEL UNIVERSITY	2	a) 7-4-1 b) -7+8	Data not available	
7. I.M.A. NUMERACY 1977		a) 77-53 b) 911-102		

TABLE:4 DIFFICULTIES AND ERRORS IN MULTIPLYING WHOLE NUMBERS

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I	5	From 38 to 567 x9 x412	Separate column multiplication (9 yr. old is 3D by 3D) Regrouping or sub-totalling (13 yr, 17 yrs)	47%
2. N.A.E.P. II	6	From 32 to 671 x4 x402	Separate column multiplication as in I	up to 44%
3. BRITISH COLUMBIA	1	403 x59	Place value errors (mis-judgment)	16%
4. A.P.U. 1977-78	?	a) 76 x 7 = b) 381 x 11 = c) 314 x801	Data not so far available	
5. SCHOOLS COUNCIL	2	a) 392 x7 b) 283 x7	Errors in basic facts Errors in carrying	
6. BRUNEL UNIVERSITY	-			
7. I.M.A. NUMERACY 1977	2	a) 7 x 8 = b) 6 x 79=	Data not so far available	

TABLE:5 DIFFICULTIES AND ERRORS IN DIVIDING WHOLE NUMBERS

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I 1972-73	3	a) 5)$\overline{125}$ b) 5D÷1D R=0 c) 3D÷2D R=0	I don't know (9 yr) No data so far available	38%
2. N.A.E.P. II	4	a) as in a) above b) 2D÷1D R=0 c) 6)$\overline{608}$ d) 28)$\overline{3052}$	No data so far available	
3. BRITISH COLUMBIA 1977	1	45)$\overline{1232}$	5x4=30 9x4=26 Place value	9% 9% 4%
4. A.P.U. 1977-78	?	816 divided by 8 (format not known)	Data not so far available	
5. SCHOOLS COUNCIL	2	a) 255÷6 b) 315÷6	20% did not attempt either Other data not available	
6. BRUNEL UNIVERSITY	-			
7. I.M.A. NUMERACY 1977	2	a) 24÷6 b) 243÷9	Data not so far available	

TABLE:6 DIFFICULTIES AND ERRORS IN ADDING COMMON FRACTIONS AND MIXED NUMBERS

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I 1972-73	1	1/2+1/3=	Wrong algorithm a) add both numerators and denominators (2/5) b) add only denominators (1/5)	16-30% 6-9%
2. N.A.E.P. II 1977-78	7	From 4/12+3/12= via 1/2+1/3= to 41/4 +32/5	add both numerator and denominator Raw Data Not Examined	11-24%
3. BRITISH COLUMBIA 1977	1	1/2+1/3	a)add both numerator and denominator b)add only denominator	9-19% 2- 8%
4. A.P.U. 1977-78	?	1/16+1/2= 1/6 +2/3=	Data not so far available	
5. SCHOOLS COUNCIL	-			
6. BRUNEL UNIVERSITY	1	9/16+5/64=	Data not available	
7. I.M.A. NUMERACY	-			

TABLE:7 DIFFICULTIES AND ERRORS IN SUBTRACTING COMMON FRACTIONS AND MIXED NUMBERS

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I 1972-73		No items included		
2. N.A.E.P. II	3	a) Type 3/5-1/5= b) Mixed numbers no common denominator, no borrowing c) as b) with borrowing	Similar to addition e.g. subtract denominator	
3. BRITISH COLUMBIA 1977	1	2/5-1/3=	a) subtract both numerator and denominator (1/2) b) multiply denom. cross multiply and add numerator (11/15)	7-22% 5-9%
4. A.P.U. 1978	?	Fractions Fractions with different denominators		
5. SCHOOLS COUNCIL	-	No pure computation items included		
6. BRUNEL UNIVERSITY	-	No pure computation items included		
7. I.M.A. NUMERACY	-	No pure computation items included		

TABLE:8 DIFFICULTIES AND ERRORS IN MULTIPLYING COMMON FRACTIONS AND MIXED NUMBERS

PROJECT	DESCRIPTIONS OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I	2	a) 1/2x1/4= b) Fraction x whole number	Incorrect algorithms of several kinds	11-15%
2. N.A.E.P. II	6	3/4x3/5(also 3 similar)	Divided 1st by 2nd	3-7%
		31/2x62/3	Separate multiplication of whole numbers and fractions	14-27%
		1/6x71/3	Inability to reduce fractions correctly	
3. BRITISH COLUMBIA	1	3/4x5/7	Invert and multiply a) 1st by 2nd b) 2nd by 1st	10-12%
			Add both numerator and denominator	1%
4. A.P.U.	?	Two simple fractions	No data so far available	
5. SCHOOLS COUNCIL	-			
6. BRUNEL UNIVERSITY	-			
7. I.M.A. NUMERACY	-			

TABLE 9: DIFFICULTIES AND ERRORS IN COMPUTATIONS WITH DECIMALS

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
1. N.A.E.P. I 1972-73	1	Addition ($and¢) Vertical format	Decimal error Carrying error	2%-22% 2%-5%
	1	Subtraction Horizontal word format		Decimal place error 4% Regrouping 3% Subtraction reversal 2%-6%
	1	Multiplication (Type .8x5.6)	Misplaced decimal point	16%-26%
2. N.A.E.P.	7	ADDITION $.70+$.40+$.20 .70+ .40+ .20	Performance level markedly lower with "no money" content	
		.35 +1.4 .35 + 1.4	Horizontal format results in ignoring fractional place value e.g. .49	Approximately half of wrong answers in this category
		7.54 1.52 7.54 + 1.52 3.57 + 1.2	Similar horizontal efforts Wrong placement of decimal Errors greater when number of fractional digits differs	3%-34%
	7	SUBTRACTION	Similar errors of misalignment and horizontal format	
	4	MULTIPLICATION 4 items involving 1 or 2 digit factors	Alignment not important, main errors misplaced decimal	

TABLE 9 (Cont't)

PROJECT	DESCRIPTION OF ITEMS		ERRORS	
	NUMBER	EXTENT	NATURE	EXTENT
	3	DIVISION Type 3)8.7 Type 8.7÷3	Main error misplaced decimal point	
3. BRITISH COLUMBIA	1	Addition ($+¢) Vertical format	Carrying error Misalignment of decimal point	8% 3%
4. A.P.U.	4	MULTIPLICATION What number is 10 times 0.5? How many times is 0.1 greater than 0.01? 12x2.54 0.5x6	No data so far available	
	2	DIVISION 1.8÷3 40÷0.8	No data so far available	
5. SCHOOLS COUNCIL	2	SUBTRACTION 3.0-0.4=	Not attempted Misuse of decimal point a)0.1 b)1.0 or 1 Adding (3.4)	21% 9% 6% 9%
		What number do you get if you take 0.4 away from 3.0?	Similar to above Not attempted 1.0 or 1 0.1 Adding	 14% 9% 9% 4%
6. BRUNEL UNIVERSITY	2	MULTIPLICATION 210 x 3.5 is 0.3 x 0.3 is	No data available Misalignment of decimal point 0.9	 c.f. 0.4x 0.4=0.16
	1	DIVISION 8:0.16 is	No data available	
7. I.M.A. NUMERACY	1 1 1 1	13.3+2.8 79.3-8.1 3x42.5 13.5÷5	No data available	

The Relation Between Elementary Algebra Students' Errors in Arithmetic and Algebra in Selected Types of Problems

Dr. Mary Ann DeVincenzo
Fordham University, NYC

Although students' computational mistakes are sometimes judged to result from carelessness and a lack of knowledge of basic number facts (Ashlock, 1976, p. 7; Easley, 1977, p. 7), it has been demonstrated that in certain instances incorrect mathematical solutions are accompanied by faulty solution techniques. For example, the application of the faulty algorithm of adding numerators and adding denominators to solve the problem 1/2 +1/3 results in a solution of $\frac{1+1}{2+3} = \frac{2}{5}$. This study considers those errors which are judged to result from the application of a faulty algorithm.

It has been observed that some students: (a) apply a faulty algorithm to solve several problems of the same type (Ashlock, 1976; Cox, 1974; Graeber & Wallace, 1977; West, 1971), (b) continue to apply the same faulty algorithm to an arithmetic problem one year later (Cox, 1974), and (c) apply the same error pattern to arithmetic problems of different degrees of difficulty which require similar skills (Graeber & Wallace, 1977). Some of the same error patterns have also been reported from groups of students' mathematics solutions in different grades (Cox, 1974; Mackay,1977). Therefore, the purpose of this study was to investigate how errors in arithmetic, identified as being persistent in nature, were related to elementary algebra students' errors in algebra for eight types of problems.

More specifically, the study was designed (a) to determine if elementary algebra students tend to make errors in algebra which are essentially the same as their errors in arithmetic, and (b) to probe and gain insights into the nature of specific errors in arithmetic and algebra.

Instrumentation

A diagnostic test in arithmetic and algebra was constructed as an instrument to help investigate relationships between errors in arithmetic and algebra. The arithmetic and algebra parts of the test each contained three examples for each of the eight types of problems selected for investigation to allow for consistent errors. A consistent error was defined as the incorrect solution of at least two problems of a specific type within one form (arithmetic or algebra) which resulted from the application of one specific faulty algorithm. Content validity and the similarity of corresponding problems on the arithmetic and algebra parts of the test were established by a panel of experts. The coefficient of internal consistency was computed to be .8624 for the arithmetic test and .8507 for the algebra test from the Kuder-Richardson 20 Formula.

The Observation-Interview session was utilized as an instrument to: (a) help establish the reliability of identifying specific errors from students' written solutions on the diagnostic test and (b) help define the nature of specific errors that were identified from written solutions on the diagnostic test. In a typical Observation-Interview session the student was observed "thinking aloud" as he/she solved a problem and questioned about specific parts of a solution which needed to be clarified. For example, an Observation-Interview session with Student 807 as he solves 5/8 - 1/3 is reported.

> I was thinking of a denominator that these two would both go into. I'm going to try two ways of doing this. That's what I usually do. I keep on going until I find something that suits me. I was thinking of timing these. 15 minus 8, but I think the answer would be in a fraction. So if I didn't do it this way, if I thought of another way. I looked at the 3 and 3. So I said let's see what they both go into. So I was thinking of timing it, timing it on both sides. I have 5 over 1 left. So I would have said 5 times 3 is 15. Then minus 8 is 7. But since this is a fraction. Since they're both fractions subtraction, I thought maybe multiplication would apply to fractions. So I tried it out. It didn't look right. I decided to just subtract 5 and 1 and the 8 minus 3. So that's what I did. I thought maybe the answer would be in a fraction.

Why do you think the answer might come out to be a fraction?

> Because this is a subtraction problem and it's in fractional form. And so I thought I probably could use multiplication. Since it's in fractional form, I'd use the opposite of division. Since they're all fractions I was thinking of multiplication.

How did you decide to subtract 1 from 5 and 3 from 8?

> Well, I thought the only way I would get a fraction, since I saw I wasn't satisfied with the answer. So I decided to get it all in a fractional form. So this is how I did it. This is the quick. To me it was the best, the fastest way to get it, a fraction over a fraction.

Procedure

During September, 1978, 1,517 ninth grade elementary algebra students in 14 public and parochial schools in 5 boroughs in New York City were given the arithmetic part of the diagnostic test. The subjects can be described as students who generally tested on ninth grade to two years below ninth grade level in mathematics. The algebra part of the diagnostic test was administered during the Spring, 1979 semester when all the topics on the algebra test were taught to those students who continued in the elementary algebra sequence. There were 1,212 students who were given the algebra test with only 1,122 of these students who were administered both the arithmetic and algebra parts of the diagnostic test.

On the basis of a pilot study conducted during the spring 1978 semester, procedures were established for coding and recording students' solutions to the diagnostic test. The coding system allowed for numbers and types of errors to be compiled by computer for each problem type on each part of the test. This coding system also permitted a comparison of solution techniques that were used for corresponding problem types on the arithmetic and algebra parts of the diagnostic test.

Observation-Interview sessions were conducted with 93 students whose solutions showed indications of consistent errors. A same error was defined to be the same consistent error and a different error was defined to be different consistent errors used to solve corresponding problem types in arithmetic and algebra. Taped Observation-Interview sessions were conducted approximately two weeks after the administration of the algebra test in each of the 14 schools. These taped sessions were transcribed and highlights recorded on an Individual Observation-Interview session data sheet for each student who participated in an Observation-Interview session.

Data Analysis and Results

The tendency for students to make the same error in arithmetic and algebra was found to vary with the type of problem. Students tended to make the same error in arithmetic and algebra for addition and subtraction of fraction problems and different errors in arithmetic and algebra for multiplication, simplification and division of exponential expressions problems.

Using a chi-square test for one-dimensional arrays and a significance level of $p<.05$, it was found that:

1. There were significantly more students who made the same error than different errors for arithmetic and algebra in addition and subtraction of fraction problems.

2. There were significantly more students who made different errors than the same error for arithmetic and algebra multiplication, simplification and division of exponential expression problems.

3. There was no significant difference between numbers of students who made the same error and different errors for arithmetic and algebra multiplicaton of fraction problems.

4. There was a significant difference in the types of errors that were repeated in arithmetic and algebra for addition, subtraction and multiplication of fraction problems. The most frequently occurring errors that were repeated in arithmetic and algebra were:

(a) adding numerators and adding denominators ($\frac{1}{2} + \frac{1}{3} = \frac{1+1}{2+3} = \frac{2}{5}$),

(b) subtracting numerators and subtracting denominators ($\frac{5}{8} - \frac{1}{3} = \frac{5-1}{8-3} = \frac{4}{5}$), and

(c) finding equivalent fractions with the same denominator and multiplying the numerators and using the common denominator for the answer ($\frac{4}{5} \times \frac{2}{3} = \frac{12}{15} \times \frac{10}{15} = \frac{120}{15}$) for addition, subtraction and multiplication of fractions problems respectively.

5. There were significantly more instances in which students made a consistent error in only algebra than in both arithmetic and algebra for problems which involve the Distributive Principle in multiplicaton and division.

6. There were significantly more students who used different algorithms than the same algorithm to solve arithmetic and algebra problems which involve the Distributive Principle in multiplication and division when a consistent error occurred in only algebra for these problems. For example, students who express $2(x+3+y)$ by $2x+3+y$ tended to compute $5(2+3+4)$ by $5(9) = 45$.

Insights into the nature of 41 specific error types were obtained from an analysis of the taped Observation-Interview session. For example, interviews revealed that some students who added numerators and added denominators to solve addition of fraction problems: (a) were guided by the position of the "+" sign in the middle of the two fractions, (b) added across because you can add whole numbers by adding across, (c) based this solution on going across to solve multiplication of fractions problems, (d) experienced difficulties with finding a common denominator and applying the standard algorithm before adding across, and (e) added across from left to right because they read from left to right.

Discussion

The tendency for students to make the same error in arithmetic and algebra for addition and subtraction of fractions problems implies that remediation in arithmetic will most probably also produce "correct" solutions for algebra addition and subtraction of fractions problems. However, since students tended to apply algebraic techniques to arithmetic as well as algebra problems during the interviews, it seems that learning how to solve correctly either the arithmetic or algebra problems will also produce "correct" solutions for both arithmetic and algebra addition and subtraction of fractions problems. For example, one student who computed $7^6 \text{ X } 7^5$ by multiplying 7 eleven times on the diagnostic test in September expressed the answer as $(7X7)^{6x5}$ during the interview and explained: "The only reason I could think of is because I find it easier to do it like that. Cause if you do it the other way it will take you more time to figure out the answer."

The tendency for students to make different errors in arithmetic and algebra for multiplication, simplification and division of exponential expressions problems and errors in only algebra for multiplicaton and division problems which involve the distributive principle was observed from results of the diagnostic test. The Observation-Interview sessions also highlighted these tendencies by revealing that students perceived the corresponding arithmetic and algebra problems differently for these 5 problem types and associated a specific algorithm with each form of the problem. For example, a student multiplies the exponents to express $b^3 \times b^2$ as b^6, attributing this solution to the x sign and explains, "If there were numbers then I'd do it out." It is true that for all 5 types of problems there are alternative techniques available to solve the arithmetic problems. However, such alternatives cannot always be used to solve the analogous algebra problems. It is important therefore, that students become familiar with alternative techniques in arithmetic that will be needed in algebra.

Learning either the arithmetic or algebra solutions for specific types of problems necessitates the creation of appropriate tasks designed to meet the needs of individual students. Since many faulty algorithms were based upon erroneous logic, this implies that there is a logical system for which "correct" algorithms can be taught and also learned. Also, in many instances specific faulty algorithms were based upon solution techniques intended to solve other types of problems. Some examples of these error types are: (a) for subtraction of fractions problems, multiply each fraction by a common denominator ($\frac{5}{8} - \frac{1}{3} = (24)\frac{5}{8} - (24)\frac{1}{3} = 15-8=7$) originally taught to solve fractional equations, and (b) for multiplication of fractions problems, invert the second fraction and multiply ($1/2 \times 1/4 = 1/2 \times 4/1 = 4/2 = 2$), which is the standard algorithm taught for dividing fractions; express each fraction as an equivalent fraction with the same denominator and multiply numerators and use the common denominator in the answer ($1/3 \times 2/9 = 3/9 \times 2/9 = 6/9$), an adaptation of the addition of fractions algorithm; multiply original numerators and use common denominator as denominator in answer ($4/5 \times 2/3 = 4/15 \times 2/15 = 8/15$) an adaptation of the algorithm to add fractions, and multiply each numerator with denominator of other fraction and place answer in either numerator or denominator ($1/2 \times 1/4 = 4/2$ or $2/4$), taught as cross-multiplication to solve fractional equations.

Consequently, a specific technique presented as a solution for one type of problem may be an important factor in a student's consideration of utilizing this technique to solve a different type of problem. An example of this was observed in this study. In one particular school students were not taught to solve fractional equations by the standard algorithm of multiplying both sides of an equation by a common denominator. Therefore, no instances of multiplying by a common denominator in addition, subtraction or multiplication of fractions problems were recorded for this school. Future research can consider presentations of standard algorithms being taught and alternate solution techniques for specific standard algorithms that may not possess the same tendency for use as solutions for other types of problems.

Analysis of the solution techniques used for addition, subtraction and multiplication of fractions problems on the diagnostic test and during the interview sessions suggested that a possible relationship may exist among the solutions to these three types of problems. Students who tended to "go across" to solve addition and subtraction of fractions problems ($\frac{1}{3} + \frac{2}{7} = \frac{1+2}{3+7} = \frac{3}{10}$ and $\frac{7}{9} - \frac{4}{5} = \frac{7-4}{9-5} = \frac{3}{4}$) did the same for multiplication of fractions problems (1/5 x 2/9 = 2/45)resulting in a correct solution for multiplicaton and incorrect solutions for addition and subtraction of fractions. Another tendency observed was that students who applied the standard algorithm taught for addition and subtraction of fractions in these problems also applied an adaptation of this algorithm for multiplication of fraction problems. For example, computing 1/2 X 1/4 by this method results in a solution of 2/4 X 1/4 = 2/4. These conjectured relationships between solutions to addition, subtraction and multiplication of fractions problems suggest that future research investigate such relationships.

Results of the Observation-Interview sessions revealed that in certain instances students attempted to apply the standard algorithm to solve addition and subtraction of fractions problems, although unsuccessfully, before resorting to "going across". Illustrations of this are contained in the following excerpts:

> You have to get a number that goes into these two (denominators) and then you put it down here with b and d. I don't know how to do the rest . . . Just add it like this, bring it over. (Student 957)

> I'm trying to find out the common denominator. I tried to find a common denominator but I couldn't find it . . . It's a simpler way (adding across) without having to do all that work. (Student 1286)

Future research could investigate the extent of the tendency to "go across" after an initial attempt with the standard algorithm if the orignial statement of the problem is presented in a vertical format rather than a horizontal format.

The role of the Observation-Interview session and its importance for teaching in general was clearly exhibited in this study. The Observation-Interview sessions provided the researcher with more detailed information that helped written solutions on the diagnostic test be better understood.

Certain written solutions appeared to represent "correct" applications of standard algorithms. However, the Observation-Interview sessions discovered that these seemingly "correct" solutions were accompanied by erroneous mathematical reasoning.

Some students performed the computations correctly in multiplication of fraction problems because:

> They're both numbers and they're both letters. ($4/a \times 2/b = 8/ab$) (Student 1504)
>
> This is the right way because you put the letters together. ($4/a \times 2/b = 8/ab$) (Student 36)
>
> That's straight across like that . . . Cause it's in a direct pattern straight across the page. ($1/3 \times 2/9 = 2/27$) (Student 179)
>
> Probably believe the way the numbers are arranged. You just do one set of numbers at once. ($1/3 \times 2/9 = 2/27$) (Student 983)

In Simplification of Exponential Expressions problems ($(a^2)^3$) some students arrived at a correct solution because:

> They're in parentheses, you multiply. (Student 885)
>
> Parentheses and the number outside indicated to multiply. (Student 589)
>
> The number all the way on the top is much smaller than the first number. So I thought if I multiplied both of them since they're both small and this one's outside the bracket, I'd see what I get since they're both sort of different from th 2 (base) here. (Student 807)

Certain other solution techniques were incorrectly identified and described by looking at only the written solutions. For example, in addition of fractions problems, many students were observed to compute $a/b + c/d$ as ac/bd . It seems as if this solution respresents the multiplication of numerators and multiplicaton of denominators. However, the interviews clarified that students were actually adding numerators and denominators. This is verified by several excerpts:

> a plus c is ac and b plus d is bd . . . You don't know what they are so you'd have to substitute for them. So you'd just put ac . . . If you had two numbers, let's say this was 2 and this was 4, you'd add them because you have to combine them cause that's what adding means. So you figure that if you had a and c instead of numbers, you combine them. So combining them is like putting it together. It's a and c together. (Student 1319)
>
> Just added them together . . . I already put them together. Like I'll use an example. If I break this pencil, I'll have one half and one part and after I put it together I won't have no more half. I'll have a whole pencil. So the a and the c when you put them together you have no more a plus c. You have ac and the same with bd. (Student 1181)

In another instance during the Observation-Interview sessions, the omission of the plus (+) sign in the solution to an addition of fractions problems was reasoned to be similar to the omission of the times sign (x) in multiplication problems, "Because in multiplication if there was an x here you wouldn't put the x there. That's why I didn't put the plus there." (Student 973)

In most instances "correct" as well as incorrect solution techniques were applied by rote during the Observation-Interview sessions. Many erroneous techniques were found to be based upon techniques taught as solutions for other types of problems. This suggests that the relationship of arithmetic and algebra errors to teaching methods and students' potential for learning mathematics be explored and compared.

Annotated Bibliography

Alexander, D.W. A clinical diagnostic investigation of fundamental reasons for conceptual difficulties with algebra (Doctoral dissertation, State University of New York at Buffalo, 1977). *Dissertaton Abstracts International,* 1978, *39*, 764A. (University Microfilms No. 78-13, 995)

This study analyzes interviews conducted with 12 ninth grade elementary algebra students within a cognitive model which combines Bruner's and Erlwanger's models.

Ashlock, R.B. *Error patterns in computation: A semi-programmed approach*. Cincinnati, Ohio: Charles E. Merrill Publishing Company, 1976.

Samples of students' incorrect solution techniques for operations with whole numbers, decimals and fractions are presented. These examples illustrate specific error patterns and possible remedial activities are also discussed.

Bachman, C.A. Analyzing children's work procedures. In M.N. Suydam (Ed.), *The 1978 National Council of Teachers of Mathematics Yearbook: Developing Computational Skills*. Reston, VA: The National Council of Teachers of Mathematics, 1978.

Eight categories of student work procedures and answer patterns are discussed.

Bright, G.W. Assessing the development of computation skills. In M.N. Suydam (Ed.), *The 1978 National Council of Teachers of Mathematics Yearbook: Developing Computational Skills*. Reston, VA: The National Council of Teachers of

Mathematics, 1978.

An analysis of the 1972 National Assessment of Educational Progress (NAEP) tests and the 1968 National Longitudinal Study of Mathematics Abilities (NLSMA) test results identifies areas which present mathematical difficulties for students.

Callahan, L.G. Clinical evaluaton and the classroom teacher. Paper presented at the AERA meeting, New Orleans, February 25-March1, 1973. (ERIC Document Reproduction Service No. ED 134 468)

The vaue of the interview for the classroom teacher is discussed as a means of evaluating students.

Carepenter, T., Coburn, T., Reys, R. & Wilson, J. Notes from natonal assessment: Addition and multiplication with fractions. The Arithmetic Teacher, 1976, 23, 137-141.

Analysis of students' solutions to several problems on the 1972 NAEP tests indicates students' level of proficiency in adding and multiplying fractions.

Cleary, M.F. A study to investigate areas of algebraic deficiency among college students and the effectiveness of a program of remediation concurrent with a beginning college mathematics course (Doctoral dissertation, University of Maryland, 1976). Dissertation Abstracts International, 1976, 36, 7006A. (University Microfilms No. 77-9, 496).

The effectiveness of a remedial college mathematical program and several common error patterns in operations with exponential expressions in algebra are presented.

Cox, L.S. Analysis, classification and frequency of systematic error computational patterns in the addition, subtraction, multiplication and division vertical algorithms for grades 2-6 and Special Education Classes. Kansas City, Kas.: Kansas University, 1974. (ERIC Document Reproduction Service No. ED 092 407)

A discription of systematic error patterns in the 4 whole number vertical algorithms for grades 2-6 is presented for an entire academic year.

Crouse, R.J. & Sloyer, C.W. Mathematical questions from the classroom. Boston: Prindle, Weber & Schmidt, Inc., 1977.

Examples of faulty algorithms used to solve various arithmetic and algebra problems are illustrated.

Easley, J.A., Jr. On clinical studies in mathematic education. Champaign, Ill: University of Illinois at Urbana-Champaign, 1977.

The value of interviews to the researcher and classroom teacher is discussed using examples of clinical research studies.

Feghali, I. Interviews with students of high confidence and low achievement. Vancouver, B.C.: Mathematics Education Diagnostic and Instructional Centre, University of British Columbia, 1976. (ERIC Document Reproduction Service No. ED 128 230)

Interviews with students making computational errors demonstrate that some students have high confidence in their faulty algorithmic techniques.

Francks, M. An analysis of errors in ninth grade algebra and the use of this knowledge in instrctional plans. Unpublished doctoral dissertation, Columbia University, 1953.

A list of algebra errors and percentages of students making errors on written tests for an academic year are presented for 1000 ninth grade elementary algebra students.

Gattegno, C. The use of mistakes in the teaching of mathematics. The Mathematics Gazette, 1954, 37(323), 11-14.

Probing into the nature of algebra errors is presented as a way of improving the teaching and learning of algebra.

Graeber, A.O. & Wallace, L. Identification of systematic errors: final report. Philadelphia: Reasearch for Better Schools, Inc., 1977. (ERIC Document Reproduction Service No. ED 139 662)

Faulty algorithms used by students in grades 1-6 in 2 schools during the 1974-75 school year are reported from completed Individually Prescribed Instruction (IPI) tests.

Kent, D. The dynamic of put. Mathematics Teaching, 1978, 82, 32-36.

Students' explanations of erroneous algebraic techniques are discussed.

Lankford, F. Some computational strategies of seventh grade pupils, final report. Charlottesville, Va.: The University of Virginia, 1972. (ERIC Document Reproduction Service No. ED 069 496)

Descriptions of error patterns for the four operations with fractions obtained from oral interviews with seventh graders are reported.

Laursen, K.W. Errors in first year algebra. The Mathematics Teacher, 1978, 71, 194-195.

Some common errors in elementary algebra are listed.

MacKay, I.D. *A comparison of students' achievement in arithmetic with their algorithmic confidence*. Vancouver, B.C.: Mathematics Education Diagnostic and Instructional Centre, University of British Columbia, 1975. (ERIC Document Reproduction Service No. ED 128 228)

Many students in grades 5-8 are reported to have high algorithmic confidence in their erroneous solution techniques for problems on the 4 operations.

Shevarev, P.A. (An experiment in the psychological analysis of algebraic errors.) In J.W. Wilson (Ed.) & A. Leong (Trans.), *Soviet studies in psychology of learning and teaching mathematics* (Volume 12). Chicago: University of Chicago Press, 1975. (Reprinted from *Proceedings of the Academy of Pedagogical Sciences of the RSFSR*, 1946, *3*, 135-180)

An analysis of specific algebra errors is presented with possible explanations for the errors.

Smith, A. An attempt to eliminate common errors in elementary algebra. *The Mathematics Teacher*, 1946, *39*, 385-390.

Possible reasons for errors with the distributive property in algebra problems are hypothesized.

West, T.A. Diagnosing pupils' errors: Looking for patterns. *The Arithmetic Teacher*, 1971, *18*, 467-469.

The importance of searching for patterns in students' incorrect computational work is stressed.

Diagnosing Understandings of Equivalences

George W. Bright
University of Wisconsin-Madison
Northern Illinois University

John G. Harvey
University of Wisconsin-Madison

Understanding equivalences seems to be important for learning mathematics at almost every level. For example, young students learn that 5 + 2 and 7 are equivalent ($5 + 2 = 7$); older students learn that 5/6 and 10/12 are equivalent ($5/6 = 10/12$); and still older students learn that the opposite of the opposite of 9 is 9 ($-(-9) = 9$). Mathematically-oriented students eventually learn that $3x + 2 = 7$ and $3x = 5$ are equivalent, but how the average algebra student perceives this relationship is unclear. If asked, students may say an equivalence ("3/4 is 75 hundredths"), but they may not perceive the relationship between this utterance and the correct written response of "0.75" to the exercise, "Write a decimal equivalent to 3/4."

As important as the concept of equivalence is to mathematics learning, there appears to be only a little literature which reports direct investigation of students' perceptions about it. In order to focus on equivalences, therefore, it was necessary to examine literature which indirectly pertains to this topic. Two instances of equivalence topics were selected from the school mathematics curriculum and the literature was searched for information on the description of student errors for these parts of mathematics. The two instances of equivalences selected were equivalences between common fractions, decimal fractions and percents, and equivalences between linear equations, especially in the context of solving linear equations.

These two instances of equivalence are related in that both deal with rational numbers. In the first case, direct manipulation of rational numbers is involved, and in the second case, the typical axiomatic approach toward equation solving involves generating equivalent equations through operations with rational numbers. Comparison of two embodiments of equivalences provides the possibility of deeper analysis than is possible with only one embodiment.

Much of the literature dealing with common and decimal fractions, with percents, and with solving linear equations is categorizable as methods-comparison studies which report only measures of overall performance (e.g., McMahon, 1960; Davis, 1976; Settle, 1977). Other studies have provided status data on difficulty of items or on specific solution strategies (e.g., Montgomery, 1959; Weaver, 1971; Brandner, 1976). Such information may be useful in improving instruction, but it is not very helpful in understanding student perceptions of equivalence. These studies are, therefore, not discussed.

For each instance of equivalence, student errors that have been identified in the literature fall into three classes: computational, non-categorizable, and potential misunderstandings of equivalence. Purely computational errors are of no interest for this paper. For example, in the student-produced sequence below the student replaced 6-10 with 4 instead of -4.

$$2x + 4 = 6 - 10 + 4x$$
$$2x + 4 = 4 + 4x$$

This error is purely computational and does not suggest any misunderstanding of equivalence of linear equations. Computational errors will not be pursued in this paper.

Non-categorizable errors are errors about which too little is known to permit interpretation. These errors may be caused by misunderstandings about equivalences, but there is not enough information available to hypothesize a cause for the error. Usually, researchers do not report the specific errors, but rather report only the number of non-categorizable errors.

Potential misunderstandings of equivalence, however, are of interest. In the sequence,

$$8/15 = \underline{\hspace{3cm}}$$
$$8/15 = \underline{\quad 15 \div 8 \quad}$$
$$8/15 = \underline{\quad 1\frac{7}{8} \quad}$$

there is a possibility that the student does not understand when a fraction and a mixed number are equivalent. These errors will be the primary focus of this paper.

Equivalence of Fractions and Percents

Rational numbers may be represented in three ways: (a) as quotients of integers (referred to as common fractions in this paper), (b) as terminating or repeating decimal fractions (referred to as decimal fractions or decimals in this paper), or (c) as percents. Students learn several different transformations among these representations; the transformations discussed in this section are from

1) a common fraction into an equivalent common fraction,

2) a common fraction into an equivalent decimal fraction or a decimal fraction which approximates, to the desired number of decimal places, the given common fraction,

3) a common fraction into an equivalent percent or into a percent which approximates, to the desired number of

decimal places, the given common fraction,

4) a decimal fraction into an equivalent common fraction,

5) a decimal fraction into an equivalent percent

6) a percent into an equivalent common fraction, and

7) a percent into an equivalent decimal fraction.

The first transformation is important for the usual algorithms for adding or subtracting common fractions, reducing common fractions, or changing between an improper fraction and a mixed number. The remainder are useful in many situations; for example, to compare value when one store advertises "1/3 off" and another "Marked down 30%." Thus, students would seem to benefit from acquisition of skill with these transformations.

A large minority of our school population has not acquired these skills. The 1973 and 1978 mathematics assessments by the National Assessment of Educational Progress (NAEP) provide some evidence of this (NAEP, 1977 and 1979). In their discussion of the 1973 NAEP results, Carpenter, Coburn, Reys, and Wilson (1978) concluded for 13- and 17-year-olds and for young adults, that expressing a given number in a different form appeared to be a difficult task and that the basic notions of percent were generally weak. Based on examination of the 1978 NAEP results, Carpenter, Corbitt, Kepner, Lindquist, and Reys (in press) reported that about two-thirds of the 13-year-olds and three-fourths of the 17-year-olds could reduce fractions to lowest terms, find equivalent fractions having different denominators, and correctly make the mixed number to improper fraction transformation; that about 50% of both age groups could change decimal fractions into common fractions; and that about 33% of the 13-year-olds and 50% of the 17-year-olds could transform a fraction into a percent. To understand why students do not perform better with these equivalences it is important to discover what errors students are making in applying the seven transformations described earlier.

The existing data about errors are scant. Apparently, there is no research at all on the kinds of errors students make when applying the decimal to common fraction transformation. Data on errors made when the other six transformations are used come from two sources: research studies conducted before 1950 and the second (1978) NAEP mathematics assessment. Thus, the conclusions reached will be severely limited. Any conclusions derived from the "old" data will be limited by the fact that school mathematics curricula have changed since 1950 and that, therefore, students' performance and the errors which they make have changed. Conclusions based on the 1978 NAEP data will be limited because only the data for the released items could be examined; in most cases there are only a few items dealing with a specific kind of transformation. Too, in the NAEP data only the frequencies of responses are reported; errors are not reported. Thus, student errors can only be hypothesized in a post-hoc way.

Common Fraction Equivalence

There was one item, of five parts, and one item, of two parts, in the 1978 NAEP mathematics assessment dealing with common fraction equivalence. In three of the seven parts, respondents were asked to reduce a common fraction to an equivalent common fraction in lowest terms; in two parts, to produce a common fraction with a given denominator equivalent to the given one; and in two parts, to convert an improper common fraction to a mixed number. The items, the responses, and the percent of those who gave each of the responses are shown in Table 1.

See Table 1 on page 161

Reduction to lowest terms. Consideration of the classifiable responses resulted in the following list of possible errors.

1) Division of the numerator by a common factor and subtraction of the same common factor from the denominator ($3/9 = (3/3)/(9 - 3) = 1/6$).

2) Reduction of the inverted fraction ($3/9 = 3$).

3) Replacement of the numerator and denominator of the fraction by a common factor ($3/9 = 3/3$ or 1, $14/35 = 7/7$ or 1, and $18/24 = 6/6$, $2/2$, 1, or some other equivalent form).

4) Division of the numerator or denominator by the common factor and replacement of the other part by that common factor ($14/35 = 2/7$ or $7/5$; $18/24 = 4/6$; $18/24 = 3/6 = 1/2$ or something equivalent).

5) Misunderstood lowest terms ($18/24 = 9/12$ or $6/8$).

The third error was observed for all three items; the frequency of errors tended to decline with age (13-year-old versus 17-year-old data).

Equivalent fraction with a given denominator. Classifiable incorrect responses seemed to fall into four groups.

1) Added the difference of the two denominators to the given numerator and denominator ($2/3 = (2 + 3)/(3 + 3) = 5/6$; $3/8 = (3 + 24)/(8 + 24) = 27/32$).

2) Divided the quotient of the given denominators into the given numerator to find the numerator of the equivalent fraction ($2/3 = (2/2)/6 = 1/6$).

3) Replaced the numerator of the given fraction with the quotient of the given denominators (2/3 = (6 ÷ 3)/6 = 2/6; 3/8 = (32 ÷ 8)/ 32 = 4/32).

4) Multiplied the numerator of the given fraction by the denominator of that fraction (2/3 = (2·3)/6 = 6/6; 3/8 = (3·8)/32 = 24/32).

There were no released items in the 1978 NAEP mathematics assessment in which subjects were asked to reduce the given fraction to one with a smaller denominator; with such data it might be possible to delineate the error categories better. For errors two and three, the frequencies for 17-year-olds were about half those for 13-year-olds.

Improper fractions as mixed numbers. Analysis of the classified incorrect responses of the 13-year-old subjects results in the following descriptions of them.

1) Produced an equivalent fraction not in lowest terms 7/4 = 1 4/3).

2) Reduced the fraction correctly but did not produce a mixed number (14/4 = 7/2).

3) Inverted the fraction (7/3 = 3/7; 14/4 = 4/14).

Of these three errors, the most frequent was inverting the given fractions.

Conversion of Common Fractions to Decimal Equivalents

Guiler (1946a, 1946b) reported two studies dealing with this transformation. The subjects in one study (1946a) were 936 students enrolled in Grade 9 in five Ohio school systems; the subjects in the other study (Guiler, 1946b) were 925 college students who in 1938-39, 1939-40, or 1940-41, were freshmen enrolled in the School of Education at Miami University of Ohio. Each sample of students completed the Christofferson-Rush-Guiler Analytical Survey Test in Computational Arithmetic. This test consisted of six kinds of items; changing proper, common fractions to decimals, changing mixed numbers to decimals, and addition, subtraction, multiplication, and division of decimals. The number of items on this test, the number of items of each kind, and the timing of this test were not reported in either study. For each study, Guiler listed the errors which students made; many of these errors occurred with a very low frequency. Therefore, in this paper lists combining low frequency errors are presented; Table 2 reports the condensed list of errors exhibited by students when they converted proper, common fractions to decimal fractions. The three most frequent reported errors were (a) placing the denominator of the fraction on the right of its numerator and placing a decimal point between them (juxtaposing the numerator and denominator and converting the resulting whole number into a decimal), (b) dividing the denominator of the fraction by its numerator (finding the decimal equivalent of the inverse), and (c) multiplying the fraction by 100 and then correctly changing that fraction into a decimal (converting the fraction into a percent equivalent.)

See Table 2 on page 164

Table 3 is the condensed list of errors which students made when asked to convert mixed numbers to decimal fractions. The errors with highest frequency, as reported by Guiler (1946a, 1946b), were made after subjects had correctly converted the mixed number into an improper fraction and were (a) dividing the denominator by the numerator (finding the decimal equivalent of the inverse), (b) converting the numerator, denominator, or juxtaposition of the numerator and denominator of the improper fraction into a decimal fraction, and (c) multiplying the fraction or the some part of the fraction by 100 before attempting to convert the improper fraction into a decimal. It should be pointed out that in both Guiler studies and for both kinds of items, subjects often seemed to confuse the conversion of common fractions to decimals and the conversion of common fractions to percents; this is evidenced by their attempts to multiply or divide the fraction or a part of the fraction by 100.

See Table 3 on page 165

Common Fractions as Percents

In only one item in the 1978 NAEP mathematics assessment were 13-year-olds and 17-year-olds subjects asked to convert a common fraction into a percent. This item, the responses by both age groups to it, and the frequency of each response, in percent, are given in Table 4. Consideration of the incorrect responses to this item led to these hypothesized errors.

See Table 4 on page 166

1) Converted the common fraction to a decimal.

2) Read 9/100 as 9/10 and converted that fraction to a percent or decimal.

3) Divided the numerator of the fraction into the denominator.

Edwards (1930), using Compass Diagnostic Test Number XIV, tested 215 seventh-grade students. Of the 113 items on this test, 15 dealt with conversion of common fractions to percents; with decimals to percents; 15, with percents to fractions; and 15, with percents to decimals. Edwards presented a fragmentary analysis of the incorrect responses to these items and discussed only those items "to which pupils responded wrongly in the same way" (p. 630). These four groups of items will be discussed as appropriate. Five items requiring a common fraction be converted into a percent met his criterion. One of those items was the same as the NAEP item just presented except that subjects were not asked to supply the percent sign in the answer. One hundred thirty-six (63.3%) of the responses were correct; of the 79 incorrect responses, 36 (16.7%) were 90%, 17 (7.9%) were 900%, and 26 (12.1%) were scattered among 10 different responses. The incorrect response of 90% appeared among those classified in the 1978 NAEP mathematics assessment. The other, 900%, does not appear in the NAEP assessment; it may be that students multiplied the numerator of the fraction (9) by the denominator (100) and juxtaposed the percent sign.

Three of the five items presented and discussed by Edwards are conversion of improper fractions (i.e., 3/2, 9/9, and 5/4) to percents. The item responses were correct from 63.7% to 71.2% of the time. A common incorrect response to these items was a correct mixed number; this response occurred between 9.3% and 11.2% of the time. The other classified responses to 9/9 were decimals involving the whole number 99; this response is like error 2; Table 3, Guiler (1946a, 1946b).

The fifth item of this group which Edwards discussed was the conversion of 5/6. Students responded correctly to this item 119 times (55.3%). There were 36 other different responses; some responses were 62 1/2%, 86 1/3%, and 37 1/2%.

Decimal Fractions Expressed as Percents

Part of one item on the 1978 NAEP mathematics assessment dealt with this transformation. Table 5 show the item and the responses of the 13-year-olds and 17-year-olds. Only one incorrect response was classified for this part; that incorrect response indicates that subjects incorrectly placed the decimal point.

See Table 5 on page 167

Of the fifteen items pertaining to this transformation, Edwards (1930) discussed five; the decimal fractions were 0.4, 0.7, 0.0475, 0.0325, and 0.0625. From 39.1% to 48.8% of the 215 subjects responded correctly to these items. The only classified incorrect responses were those in which the decimal point was dropped from the stimulus (e.g., 4 when 0.4 was the stimulus); this incorrect response accounted for between 39.1% and 51.6% of the responses and seems akin to error 5 in Table 3 for which students, when changing a mixed number to a decimal, manipulated the mixed number to obtain a whole number and placed a decimal point somewhere in that whole number.

Percent to Common Fraction Conversion

Again, only part of one item in the 1978 NAEP mathematics assessment dealt with this transformation. Table 6 presents the data related to this item. The errors which seem to be related to the incorrect responses are described as follows.

See Table 6 on page 168

1) Made a fraction using 25 as its denominator.

2) Made a fraction using 2 as the numerator and 5 as the denominator (c.f., Error 1, Table 2).

3) Made a fraction using 25 as both the numerator and denominator.

4) Correctly changed 25% to a decimal.

Possibly, students who responded 1/5 may have first changed 25% into 1/25 and then incorrectly reduced 1/25 to 1/5.

Edwards (1930) presented data related to four of the 15 items for which students were asked to convert percents to common fractions. For each of these items, about 50% of the students responded correctly, and Edwards did not report classifications of between 36.7% and 46.5% of the responses. The incorrect responses which he did classify were (a) responding with an incorrect, widely used, common fraction (3/8 or 5/6 for 62 1/2%), (b) making a fraction using the percent given (1/6 for 6%), and (c) dividing the percent into 100 (9 1/11 for 11%).

Transforming Percents to Decimals

Part of one item in the 1978 NAEP mathematics assessment dealt with this transformation. Table 7 reports the data which were gathered. The incorrect responses suggest the following errors may have been made.

See Table 7 on page 169

1) Incorrect placement of the decimal point.

2) Made a fraction using 125 as the denominator.

3) Converted 125% into 1 1/4 and then converted 1/4 to 0.4 (125% = 1.4) (c.f., Error 1, Table 2).

4) Correctly represented 125% as a decimal and changed 0.25 into a common fraction by removing the decimal point and placing a "bar" between the 2 and 5 (c.f., Error 1, Table 2 and Error 2, Table 3).

Between 62.3% and 56.7% of Edwards' seventh-grade students responded correctly to the three items which he discussed. He classified only one incorrect response; namely, incorrect placement of the decimal point in an otherwise correct decimal. Between 16.3% and 22.8% of Edwards' subjects made this error; this was less frequent than the same error by 13-year-old students in the 1978 NAEP mathematics assessment.

This concludes the presentation of the scant data concerning the seven transformations that were described. The next section of the paper will attempt to draw some conclusions from these data.

Conclusions About Fractional and Percent Transformations

The most predominant class of difficulty seems to have been confusion in distinguishing among the three forms of rational numbers. This confusion was displayed in two ways. One way was that some students seemed to substitute one symbol for another; for example, in 7/15 = 7.15, the decimal point was substituted for the "bar." The second way was that other students applied parts of the transformations related to percent indiscriminately. Perhaps this suggests there may be interference among the concepts or among the algorithms for the various transformations. If so, the cause may be inadequately learned concepts or too rapid teaching of the concepts and algorithms.

Several less prevalent difficulties deserve to be mentioned. First, students sometimes wrote the correct value in the wrong form (e.g., 1 4/3). Second, a fraction was sometimes incorrectly inverted (e.g., 3/9 = 3). Third, a fraction was sometimes incorrectly equated to 1 (e.g., 3/9 = 3/3). Fourth, the same quantity may have been added to the numerator and denominator (e.g., 3/8 = 27/32). This last error shows up among those made by algebra students when applying operations on rational expressions (e.g., $\frac{6+x}{x+2} = \frac{6}{2}$). These four difficulties, together with the one previously discussed and two minor errors of misreading and incorrect initial operation, seem to account for all the classified errors reported.

Equivalent Linear Equations

The focus of this section is linear equations like those in a typical first-year secondary school algebra class. Although number sentences like $3 + \square = 7$ have the form of linear equations, they are not considered here since solution processes, at least for elementary school students, do not generally result in a series of equivalent equations (Carpenter & Moser, Note 1; Weaver, 1971).

Algebraic Analysis

Swain (1962) suggests two types of processes are involved in solving linear equations. One is a manipulation; e.g., adding the same thing to both sides of an equation; and the other is reduction (e.g., replacement of one expression in x by another (equivalent) expression). Romberg (1975) uses the phrase "sentential transformation" for these operations. For example,

$$\begin{aligned} 5x &= 3x + 6 & \\ 5x + (-3x) &= (3x + 6) + (-3x) & \text{manipulation} \\ 2x &= (3x + 6) + (-3x) & \text{reduction} \\ 2x &= 6 & \text{reduction} \end{aligned}$$

$1/2(2x) = 1/2(6)$ manipulation
$x = 1/2(6)$ reduction
$x = 3$ reduction

An operation transforms one equation to an equivalent equation if the solution sets and domains of the variable are the same. For example, $2x = 6$ and $x = 3$ are equivalent, but $2x = 6$ and $2 = 6/x$ are not, since 0 is in the replacement set for $2x = 6/x$ and not in that set for $2 = 6/x$. General equation solving strategies (e.g., clear fractions) may result in non-equivalent equations, but solution procedures for linear equations usually result only in equivalent linear equations. Hence, restricting attention to linear equations should simplify the search for ways to diagnose understanding of equivalence.

Errors in Solving Linear Equations

Davis and Cooney (1977) provided a reasonable categorization of the errors students made while solving linear equations. Of the five general categories of errors, two are computational and accounted for about 55% of the errors they found. The other three categories suggest possible misunderstandings about equivalent equations. These categories, and the percent of error accounted for by each category, are as follows:

1) misuse of additive or multiplicative properties of equality (13%)

2) coefficient error (9%)

3) incomplete solution after some correct steps (12%)

Eleven percent of the observed errors were unclassifiable.

Misuses of additive or multiplicative properties of equality included errors like the following:

$$3x + 5 = 28$$
$$3x + 5 - 5 = 28$$

and

$$3/4x = 28$$
$$4/3 \cdot 3/4x = 28$$

Teachers frequently note this error by saying that students have not "done the same thing to both sides of the equation." Since the resulting equation is not equivalent to the original equation, it is certainly possible that students do not understand equivalence.

Coefficient errors included errors like the following:

$$\frac{3}{2}x = \frac{5}{7}$$

$$-\frac{3}{2} + \frac{3}{2}x = \frac{5}{7} - \frac{3}{2}$$

and

$$24 = -4x$$

$$\frac{1}{4} \cdot 24 = -4x \cdot \frac{1}{4}$$

$$6 = x$$

Errors like these suggest some confusion about the operations that may be performed on an equation to produce an equivalent equation. For that reason, it seems reasonable to maintain that students who make this error may misunderstand what makes equations equivalent.

Incomplete solutions after some correct steps include student work like the following sequence. Given $4x = 7x - 36$, the student writes $4x = 7x + -36$ and then stops. These two equations together suggest that the student did not know what operations might be applied to the transposed equation. This lack of understanding may be partly due to misunderstanding why two equations are equivalent.

Students who solved 10 or 11 (of 12) equations correctly had 16% of their errors categorized in one of these three groups. Students who solved between two and seven equations correctly had 38% of their errors categorized in one of these three groups. These data suggest that investigating the understandng of equivalent equations is important in helping students who have serious difficulty solving linear equations.

It may be important in interpreting some of these errors to have measures of a student's computation skills. For example, if a student knows that 2/3 and 3/2 are inverses, but is unable to solve $2/3x = 9$, then it is reasonable to infer that the student doesn't understand that equivalent sentences can be produced by the multiplication by the inverse of the coefficient of x (3/2). This may indicate an incomplete concept of equivalence.

Similarly, if a student replaces $-2/2$ by 1, that clearly seems to be a computational error. However, if a student replaces $-2x/2$ by x, that may indicate confusion about equivalence. Of course, it may also be only a computational error. Data about students' computational skills would seem to be essential in interpreting this error.

Formal and Informal Solution Techniques

One approach to studying equation solving has been the comparison of informal and formal techniques. Informal techniques (also called cover-up techniques) are reasoning strategies whereby number relationships are used. For example, to solve $1/2 = 6/x$ students might reason that the numerators are in a ratio of 1 to 6, so the denominators must be also; that is, $x = 2 \cdot 6 = 12$.

Whitman (1976) in a study of formal and informal equation solving with 156 seventh-graders obtained follow-up data from 31 randomly selected subjects. Most of the interviewees had been exposed to informal equation solving techniques and thus, did not generate equivalent equations in typical ways. For example, to solve $3y - 5 = -(3y - 5)$, students reasoned that $3y - 5 = 0$, since 0 is the only number which is its own opposite. Then, they reasoned, "What subtract 5 is 0?"; i.e. $Z - 5 = 0$. Then, "What times 3 equals 5?"; i.e., $3w = 5$; and that is 5/3. It is not clear how many of the informally 'solved' equations were incorrectly or correctly solved. Of the 15% of the equations that were attempted by formal techniques, one-third were incorrectly or incompletely solved. There was no clear indication of the errors students made in attempting to solve these equations by formal techniques.

Petitto (1979) interviewed nine ninth-grade students using ten ratios, each of which could be solved by intuitive or formal techniques. Students generalized procedures which produce equivalent equations for solving simple ratios to more complex ratios for which equivalent equations are not produced. For example, to solve $1/2 = x/4$ students correctly multiplied by 4 to produce the equivalent equation $2 = x$; that is, the solution is the same and the domain of the variable is the same. To solve $1/2 = 2/x + 3$ students multiplied by $2(x + 3)$ to 'clear fractions' and produced the non-equivalent equation $x + 3 = 4$. Although the solution is the same, the domain of the variable is not. Students did not seem to realize that they had overgeneralized; perhaps the role of the domain of the variable had not been pointed out as being important in creating equivalent equations. The emphasis of the interviews, however, seems to have been on helping students develop intuitive approaches rather than formal ones and on obtaining solutions to the equations rather than studying or producing intermediate steps. Little importance seems to have been attached to students' understandings of equational equivalence.

Petitto speculated that the equals sign, "may mean equivalence in the sense that what is on one side is some transformation of what is on the other." (p.81) There was not clear evidence offered in support of this thesis, however. Examining understanding of equivalence between expressions may be one option in investigating understanding of equivalence between equations.

Davis, Jockusch, and McKnight (1978) extensively discussed student behavior on a variety of algebra tasks. They seem not to have dealt with equivalence of linear equations but to have focused more on problem solving and overgeneralization of rules.

Studies from Piagetian Perspective

An infrequent approach has been from a Piagetian view. Wagner (Note 2) developed a measure of conservation of equation. Two identical equations were shown and the subject was asked whether they were the same. Then the variable in one equation was changed (from $\underline{W}$ to $\underline{N}$), and the subject was asked which solution would be larger. If the response was that the solutions would be the same, the subject was said to conserve. Subjects (N = 72) at ages 12, 14, and 17 years were classified primarily as conservers (69.4%), though 22.2% were non-conservers, and 8.3% were transitional. About 50% of 12-year-olds were conservers, and about 80% of 14-year-olds and 17-year-olds were conservers. The task was easy for the subjects and seemed to focus student attention on the solutions of the equations. The study does not provide data on whether students would focus on solutions if the equations were transformed, for example, by adding the same number to both sides.

Adi (1978) used 75 prospective elementary school teachers as subjects. She taught them both formal and informal techniques and measured their formal operations levels. Subjects at the late concrete and early formal levels solved more equations than subjects at the early concrete level ($p < .01$). This result suggests another kind of information that might be useful to teachers diagnosing understandng of equivalence.

Conclusions About Linear Equation Transformations

There are few conclusions that can be drawn from such scant data. Perhaps most obvious is that students do not seem to know when equations are equivalent. This is displayed in Davis and Conney's results and in Petitto's interviews. Perhaps embedded in this confusion is some confusion about what a variable is (Tonnessen, Note 3). It is not possible to know for sure from the scanty data available.

Synthesis

Because there are no data related to students' understandings of equivalence of linear equations, it is difficult to compare the conclusions reached for rational number equivalence and sentential equivalence. The conclusions related to rational number equivalence tend to predominate, perhaps because there were more data to support probable conclusions. Nevertheless, there do seem to be some commonalities between the two sets of conclusions.

First, misunderstanding the individual objects that are to be made equivalent (i.e., common fractions, decimal fractions, percents, or linear equations) may cause difficulty in understanding whether an equivalence exists. This difficulty was evidenced when students attached an undesired percent sign in converting common fractions to decimal fractions. Too, some of the equation solving errors may be related to misunderstandings of variable which have been identified by Tonnessen (Note 3). These misunderstandngs suggest lack of a clear conception of what a linear equation represents. Firm understandng of the objects of interest in the equivalence, however, seems essential.

Second, students may not know how to choose any algorithm, or they may choose an incorrect algorithm to use in creating an equivalence. That students have to choose among several available algorithms suggests that students are working at the cognitive level of application as defined in Bloom's taxonomy. This cognitive level is above the knowledge and comprehension levels and may be higher than most educators would initially choose. In particular, application is a higher level than is required for recall or for use of a single algorithm. This difficulty of students suggests that present instruction on selecting from among different algorithms (i.e., instruction of a higher-level algorithm) may be ineffective.

Third, informal procedures used by students may interfere with formal techniques taught. The data on equation solving suggest that students more readily used informal techniques, though one can only speculate that informal or intuitive techniques may also be used in solving rational number equivalences. (For example, students may guess that 18/24 is either 2/3 or3/4 and then proceed to test these guesses.) Data would need to be generated to determine if intuitive techniques are used by students, and if so, whether they interfere with formal techniques.

Fourth, when an equivalence is not explicitly written, such as in solving linear equations. Students may not understand that an equivalence is actually called for. Mathematicians know that any two equivalence equations are implicitly joined by a biconditional "if and only if" ($\Longleftrightarrow$), but it is almost certain that students do not know this. The stress placed on memory by this implicit logical framework may cause breakdowns in processing equivalences.

Fifth, researchers who report errors made by students should give more careful analyses of those errors; too many errors are classed as "other." Too, the errors need to be more carefully classified into "sensible" categories. Guiler (1946a, 1946b), for example, created far too many categories, and as a result, the reports almost overwhelm the reader. Perhaps this means that hypotheses about errors need to be carefully checked with follow-up clinical studies. Perhaps, too, computer modeling of error behavior would help verify classification schemes.

These five generalizations have some implications for diagnosing understandings of equivalence. First, understanding of the objects of interest (e.g., fractions) should be checked. Second, questions like, "In solving $7x - 4 = 2 + 8x$, what would you do first? Second? Third?" should be part of the diagnosis procedure. Having students outline the steps in their solution without actually completing each step should yield information about how students select algorithms to apply. Too rapid focus on the solution rather than on the process of solution may indicate too little analysis of the components of the task.

Third, students' procedures should be compared with the procedures they have been taught, to determine whether informal techniques are causing interference. Fourth, students should be asked to justify critical steps in their work to determine if they perceive the logical structure of the equivalence task being solved.

On the basis of the data found, it seems safe to conclude that very little is known about students' understandings of equivalences. Given the importance of equivalence to mathematics, this lack of information is surprisng. Until data are gathered relating directly to equivalence, students' perceptions will remain in an area illuminated by half-light.

Reference Notes

1. Carpenter, T.P., & Moser, J.M. The development of addition and subtraction problem-solving skills. Paper presented at the Wingspread Conference, Racine, Wisconsin, 1979.

2. Wagner, S. Conservation of equation and function and its relationship to formal operational thought. Paper presented at the annual meeting of the American Educational Research Association, New York, 1977.

3. Tonnessen, L.H. Measuring attainment of the concept variable by college mathematics students (tentative title). Doctoral dissertation, University of Wisconsin - Madison, 1980, (in progress).

References

Adi, H. Intellectual development and reversibility of thought in equation solving. Journal for Research in Mathematics Education, 1978, 9, 204-213.

Brandner, R.J. Testing for the analytic strategy for solving linear algebraic equations (Doctoral dissertation, University of Cincinnati, (1976). Dissertation Abstracts International, 1976, 37, 164A-165A. (University Microfilms No. 76-14,544).

Carpenter, T., Coburn, T.G., Reys, R.E., & Wilson, J.W. Results from the first mathematics assessment of the national assessment of educational progress. Reston, VA: National Council of Teachers of Mathematics, 1978.

Carpenter, T., Corbitt, M.K., Kepner, H., Lindquist, M.M. & Reys, R.E. Results of the second NAEP mathematics assessment: Secondary school. Mathematics Teacher, in press.

Davis, E.J., & Cooney, T.J. Identifying errors in solving certain linear equations. The MATYC Journal, 1977, 11, 170-178.

Davis, L.H. A study of two methods of teaching problem solving in eighth grade mathematics (Doctoral dissertation, The Louisiana State University and Agricultural and Mechanical College, 1976). Dissertation Abstracts International, 1976, 37, 3373A. (University Microfilms No. 76-28, 797).

Davis, R.B., Jockusch, E., & McKnight, C. Cognitive processes in learning algebra. *Journal of Children's Mathematical Behavior*, 1978, *2* (1), 11-320.

Edwards, A. A study of errors in percentage. In G.M. Whipple (ed.), *Research in arithmetic, Twenty-ninth yearbook of the National Society for the Study of Education, Part II*. Bloomington, IL: Public School Publishing Company, 1930.

Guiler, W.S. Difficulties in decimals encountered by ninth-grade pupils. *Elementary School Journal*, 1946, *46*, 384-396. (a)

Guiler, W.S. Difficulties encountered by college freshman in decimals. *Journal of Educational Research*, 1946, *40*, 1-13. (b)

McMahon, D.L. An experimental comparison of two methods of teaching percent to seventh grade pupils (Doctoral dissertation, University of Missouri, 1959). *Dissertation Abstracts International*, 1960, *20*, 3760.

Montgomery, C.R. An investigation of the learning of the three cases of percentage in arithmetic (Doctoral dissertation, State University of Iowa, 1958). *Dissertation Abstracts International*, 1959, *19*, 1676-1677. (University Microfilms No. 58-5846).

Montgomery, J.F. *An investigation of the "case II" relationship in arithmetic*. Unpublished doctoral dissertation, Duke University, 1949.

National Assessment of Educational Progress. *Mathematics technical report: Exercise volume*. Denver, Co: The author, 1977.

National Assessment of Educational Progress. *The second assessment of mathematics, 1977-78, released exercise set*. Denver, Co: The author, May 1979.

Petitto, A. The role of formal and non-formal thinking in doing algebra. *Journal of Children's Mathematical Behavior*, 1979, *2* (2), 69-82.

Romberg, T.A. Activities basic to learning mathematics: A perspective. In *The NIE Conference on basic mathematical skills and learning*, *Volume I: Contributed position papers*. Washington, D.C.: National Institute of Education, 1975.

Settle, M.G. The relative effects of instruction in the guess and test procedure on writing relevant equations to verbal problems in algebra I (Doctoral dissertation, The Florida State University, 1977). *Dissertation Abstracts International*, 1977, *38*, 2633A-2634A. (University Microfilms No. 77-24,804).

Swain, R.L. The equation. *Mathematics Teacher*, 1962, *55*, 226-236.

Weaver, J.F. Some factors associated with pupils' performance levels on simple open addition and subtraction sentences. *Arithmetic Teacher*, 1971, *18*, 513-519.

Whitman, B.S. Intuitive equation solving skills and the effects on them of formal techniques of equation solving (Doctoral dissertation, The Florida State University, 1975). *Dissertation Abstracts International*, 1976, *36*, 5180A. (University Microfilms No. 76-2720).

Table 1

Released Common Fraction Equivalence Items from the 1978 NAEP Mathematics Assessment

Items	Responses	Percent of Total Responses*	
		Age 13	Age 17
1. Reduce the following fractions to lowest terms.			
a) $\frac{3}{9}$ = __________			
	$\frac{1}{3}$	77.8 (1.8)	92.8 (.8)
	$\frac{3}{9}$	.7 (.2)	.1 (.1)
	$\frac{1}{6}$	1.4 (.4)	.5 (.1)
	3	2.0 (.3)	1.7 (.3)
	$\frac{3}{3}$ or 1	3.1 (.4)	1.3 (.2)
	Other	8.3 (.8)	1.8 (.3)
	Don't know	1.8 (.3)	.8 (.2)
	No response	5.0 (.8)	.8 (.2)
b) $\frac{14}{35}$ = __________			
	$\frac{2}{5}$	56.9 (1.9)	77.7 (1.5)
	$\frac{14}{35}$	2.2 (.3)	2.2 (.4)
	$\frac{2}{7}$	2.9 (.4)	2.7 (.3)
	$\frac{7}{5}$	3.6 (.4)	3.1 (.4)
	$\frac{7}{7}$ or 1	1.5 (.4)	.8 (.2)
	Other	17.9 (1.3)	8.6 (.8)
	Don't know	4.0 (.5)	2.3 (.4)
	No response	10.9 (1.1)	2.5 (.4)

*The numbers in parentheses are the standard errors.

Table 1 (cont'd)

Items	Responses	Percent of total responses*	
		Age 13	Age 17
c) $\frac{18}{24} =$ ____			
	$\frac{3}{4}$	57.5 (1.7)	79.0 (1.4)
	$\frac{18}{24}$	.4 (.1)	.2 (.1)
	$\frac{9}{12}$ or $\frac{6}{8}$	4.4 (.5)	1.7 (.2)
	$\frac{1}{2}$ or $\frac{n}{2n}$	5.4 (.8)	3.3 (.4)
	$\frac{4}{6}$, $\frac{2}{3}$, or $\frac{2n}{3n}$	2.8 (.4)	3.1 (.5)
	$\frac{6}{6}$, $\frac{2}{2}$, 1, or $\frac{n}{n}$	1.9 (.4)	.8 (.2)
	Other	15.5 (1.0)	7.9 (.8)
	Don't know	2.8 (.4)	1.9 (.3)
	No response	9.2 (.9)	2.2 (.4)
2. Change each of the following fractions to an equivalent fraction have the given denominator.			
a) $\frac{2}{3} = \frac{}{6}$			
	4 or $\frac{4}{6}$	70.7 (1.6)	84.2 (1.5)
	5 or $\frac{5}{6}$	.4 (.1)	.5 (.2)
	1	12.1 (.9)	6.2 (.8)
	2	5.6 (.6)	2.3 (.3)
	6	1.1 (.3)	.3 (.1)
	Other	6.1 (.7)	3.5 (.5)
	Don't know	.7 (.2)	1.2 (.4)
	No response	3.3 (.4)	1.7 (.3)

*The numbers in parentheses are the standard errors.

Table 1 (cont'd)

Items	Responses	Percent of Total Responses*	
		Age 13	Age 17
b) $\frac{3}{8} = \frac{\quad}{32}$			
	12 or $\frac{12}{32}$	65.4 (1.5)	79.6 (1.4)
	27 or $\frac{27}{32}$	.1 (.1)	.1 (0)
	4	9.2 (.8)	4.7 (.6)
	24	1.5 (.3)	1.3 (.3)
	Other	16.6 (.9)	9.7 (.8)
	Don't know	.8 (.2)	1.6 (.4)
	No response	6.3 (.6)	3.0 (.5)
3. Write the following improper fractions as mixed numerals.			
a) $\frac{7}{3} =$ ________			
	$2\frac{1}{3}$	64.2 (1.8)	n/a**
	$1\frac{4}{3}$	.4 (.2)	n/a
	$\frac{3}{7}$	5.0 (.7)	n/a
	Other	19.4 (1.1)	n/a
	Don't know	7.8 (.9)	n/a
	No response	3.1 (.4)	n/a
b) $\frac{14}{4} =$ ________			
	$3\frac{1}{2}$ or $3\frac{2}{4}$	63.3 (1.7)	n/a
	$\frac{7}{2}$	1.3 (.2)	n/a
	$\frac{4}{14}$	4.3 (.6)	n/a
	Other	20.0 (1.0)	n/a
	Don't know	7.0 (.7)	n/a
	No response	4.1 (.5)	n/a

*The numbers in parentheses are the standard errors.
**n/a = not applicable

Table 2

Errors Made by Grade 9 Students and College Freshmen When Converting Common Fractions into Decimals

Error	Grade 9 Students*		College Freshmen+	
	Number	Percent[a]	Number	Percent[a]
1. Placed denominator alongside numerator and pointed off (7/15 = 7.15)[b]	51	5.4	21	2.3
2. Divided denominator by numerator (7/15 = 2.14)[b]	32	3.4	25	2.7
3. Multiplied fraction by 100 before changing to a decimal (7/15 x 100 = 700/15 = 46.67)[b]	29	3.1	32	3.5
4. Multiplied numerator by denominator and pointed off (7/15 = 7 x 15 = 1.05)[b]	9	1.0	3	0.3
5. Changed numerator or denominator to decimal and disregarded the other part (7/15 = 0.15)	5	0.5	0	0.0
6. Attempted to convert to a percent or treated as a percent (7/15 = 0.46%, 7/15 = 0.00 7/15, 100 ÷ 15 = 6.66, (15 ÷ 100) x 7 = 1.05, 7/15 ÷ 100 = 7/1500 = 0.0046, 7/15 = (15/7 x 100) = 214.28)	12	1.3	12	1.3
7. Unrelated errors classified by Guiler	0	0.0	4	0.4
8. Unclassified by Guiler	32	3.4	18	1.9

---

* N = 936

+ N = 925

[a] The percent was calculated by dividing the number of errors by the number of students and multiplying the result by 100.

[b] These errors were identified by Guiler (1946a, 1946b); the other errors listed combine errors listed separately by him.

Table 3

Errors Made by Grade 9 Students and College Freshmen When Converting Mixed Numbers into Decimals

Error	Grade 9 Students*		College Freshmen+	
	Number	Percent[a]	Number	Percent[a]
Changed mixed number to improper fraction and				
1. Divided denominator by numerator (6 3/11 = 69/11 = 0.159)[b]	81	8.7	134	14.5
2. Pointed off numerator, denominator, or juxtaposition of the numerator and denominator (6 3/11 = 69/11 = 0.69, 0.11, or 69.11)	22	2.4	5	0.5
3. Attempted to convert to a percent or treated as a percent (6 3/11 = 69/11 = 69/11 x 100 = 627.272, 6 3/11 = 69/11 = (100 ÷ 11) x 69 = 627.272, 6 3/11 = 69/11 = 69 ÷ 100 = 0.69, 6 3/11 = 69/11 = 69/11 ÷ 100 = 0.062, 6 3/11 = 69/11 = (11 x 100) ÷ 69 = 15.942)	9	1.0	24	2.6
4. Unrelated errors classified by Guiler	6	0.6	2	0.2
Did not change mixed number of improper fraction, but				
5. Pointed off without performing any operations (6 3/11 = 6.311, 6 3/11 = 0.06 3/11, 6 3/11 = .06, 6 3/11 = 6.3/11 = 6.11, 6 3/11 = 0.311)	45	4.9	17	1.8
6. Multiplied whole number by the denominator of the fraction and performed an incorrect operation (6 3/11 = 3 ÷ 66 = 0.045, 6 3/11 = 66 ÷ 3 = 22, 6 3/11 = 11 ÷ 66 = 0.167, 6 3/11 = 66 ÷ 100 = 0.66, 6 3/11 = 66 x 11 = 726)	14	1.5	13	1.4
7. Attempted to convert to a percent or treated as a percent	6	0.6	20	2.1
8. Incorrect initial operation (6 3/11 = 6 + (11 x 3) = 6.33, 6 3/11 = 6 + (3 + 11) = 6.14, 6 3/11 = 6 ÷ 11 = 0.545)	12	1.3	15	1.6
9. Unclassified by Guiler	39	4.2	13	1.4

* N = 936

+ N = 925

[a] The percent was calculated by dividing the number of errors by the number of students and multiplying the result by 100.

[b] This error was identified by Guiler (1946a, 1946b); the other errors listed combine errors listed separately by him.

Table 4

Released Common Fraction to Percent Conversion Item from the 1978 NAEP Mathematics Assessment

Item	Responses	Percent of Total Responses*	
		Age 13	Age 17
Express $\frac{9}{100}$ as a percent.	9%, 009%, or nine percent	34.0 (1.7)	52.3 (1.7)
	9 or 009, without label	1.6 (.3)	.4 (.1)
	.09, with or without a label	15.6 (1.0)	18.8 (1.2)
	11%, 11.1, or any attempt at $\frac{100}{9}$	1.8 (.4)	1.3 (.3)
	9 with a wrong label	.2 (.1)	.1 (0)
	90 or .9 with or without a label	11.1 (.9)	9.2 (.8)
	Other	23.0 (1.3)	10.4 (.9)
	Don't know	7.9 (.8)	4.4 (.6)
	No response	3.9 (.5)	3.1 (.5)

*The numbers in parentheses are the standard errors.

Table 5

Released Decimal Fraction to Percent Conversion Item from the 1978 NAEP Mathematics Assessment

Item	Responses	Percent of Total Responses*	
		Age 13	Age 17
Change .15 to a percent.	15% or 15	68.3 (1.5)	76.8 (1.4)
	15 with the decimal point incorrectly placed	5.2 (.7)	5.8 (.5)
	Other	11.0 (.8)	9.7 (.8)
	Don't know	8.9 (.8)	3.6 (.5)
	No response	6.5 (.7)	4.1 (.5)

*The numbers in parentheses are the standard errors.

Table 6

Released Percent to Common Fraction Conversion Item from the 1978 NAEP Mathematics Assessment

Item	Responses	Percent of Total Responses*	
		Age 13	Age 17
Change 25% to a common fraction.			
	$\frac{1}{4}$, $\frac{25}{100}$, or $\frac{n}{4n}$	56.7 (2.0)	81.2 (1.2)
	$\frac{1}{25}$	3.4 (.5)	2.9 (.4)
	$\frac{2}{5}$	4.2 (.5)	1.3 (.3)
	1, $\frac{25}{25}$, $\frac{5}{5}$, or $\frac{n}{n}$	2.8 (.5)	.6 (.2)
	$\frac{1}{5}$	1.6 (.4)	.8 (.2)
	0.25	1.8 (.4)	3.0 (.5)
	Other	15.2 (1.3)	6.1 (.6)
	Don't know	8.6 (.8)	1.9 (.4)
	No response	5.6 (.7)	2.3 (.4)

*The numbers in parentheses are the standard errors.

Table 7

Released Percent to Decimal Fraction Conversion Item from the 1978 NAEP Mathematics Assessment

Item	Responses	Percent of Total Responses*	
		Age 13	Age 17
Change 125% to a decimal fraction.			
	1.25 or equivalent decimal number	27.4 (1.5)	44.3 (1.6)
	125 with any decimal placement except 1.25	29.0 (1.1)	25.5 (1.3)
	$\frac{1}{125}$	.6 (.2)	.3 (.1)
	1.4	.6 (.2)	.6 (.2)
	$1\frac{2}{5}$	.5 (.2)	.1 (0)
	$\frac{125}{100}$, $1\frac{1}{4}$, or equivalent common fraction	6.0 (.7)	6.4 (.7)
	Other	17.7 (1.2)	12.7 (1.1)
	Don't know	11.6 (1.0)	6.2 (.7)
	No response	6.6 (.8)	3.8 (.5)

*The numbers in parentheses are the standard errors.

The Effects of Instruction and Field Experience in Diagnostic-Prescriptive Teaching of Mathematics on Preservice Elementary Teachers

Sharon L. Young
University of Colorado

Instruction for teachers in diagnostic and prescriptive mathematics has traditionally been the domain of inservice rather than preservice education. There is evidence, however, that relatively few teachers receive such training in their inservice experiences. If teachers are to learn diagnostic-prescriptive techniques, it appears that there is a need to communicate such teaching methods during preservice training.

The purpose of the study was to determine whether or not instruction and field experience in diagnostic-prescriptive teaching are viable components in a preservice mathematics methods course. An experiment was conducted at the University of Colorado, Boulder, in the mathematics methods course for prospective elementary teachers during the fall semester, 1978.

This investigation was conducted to compare the effects of instruction and field experience in diagnostic-prescriptive teaching with the effects of instruction and field experience in usual teaching methods on preservice teachers' knowledge of teaching mathematics, their knowledge and use of techniques to teach to children's mathematical strengths and weaknesses, and their attitudes towards mathematics and the teaching of mathematics. This study also compared the effects of high and low levels of preservice teachers' mathematics achievement on the same teaching and attitudinal variables.

Diagnostic-Prescriptive Techniques

In order to identify a workable combination of diagnostic and prescriptive techniques, the following aspects of diagnosis and prescription were developed for this study.

Topics for Diagnosis and Prescription

1. Content knowledge and skills
2. Levels of understanding
3. General interests and attitudes
4. Learning styles

Techniques for Diagnosis

1. Analysis of written work
2. Observation
3. Oral interviews

Techniques for Prescription

1. Setting objectives

2. Determining instructional sequence
3. Selecting appropriate materials and activities

These topics and techniques were incorporated in the experimental instruction and field experience methods that were used to test the teachability of diagnostic-prescriptive techniques to preservice teachers.

Procedure

The Setting of the Experiment

The experiment was conducted in the Elementary Professional Year Program in the School of Education at the University of Colorado, Boulder, in the fall semester, 1978. The 77 subjects of the experiment were students who were enrolled in Teacher Education (TED) 421, Methods of Elementary School Mathematics.

TED 421 is designed as a two credit preservice course for prospective elementary teachers and is offered in conjunction with other content methods courses and a required field experience course. In the field experience each student acts as an instructional assistant in an elementary school for one-half of each school day during the semester. Field assignments for the mathematics methods course require the students to teach arithmetic lessons during the field experience.

Instructional Treatments

Instruction for both treatments covered the same mathematics topics and was conveyed through the same proportion of lecture-workshops, mathematics laboratory assignments, and readings. The readings contained a common core of material, identical for both treatments, and supplemental material, different for each treatment. Two instructors each taught a control group and an experimental group.

The control instructional method stressed the use of manipulative materials and the development of mathematical concepts for teaching mathematics with understanding. References to instructional needs of individual children were not made.

The experimental instructional method also stressed teaching for understanding and emphasized the use of diagnostic-prescriptive techniques to meet the mathematics instructional needs of individual children. Diagnostic techniques included analysis of written work, observation of children, and use of individual interviews with children. Prescriptive techniques included setting objectives, determining instructional sequence, and selecting appropriate learning materials and activities. Topics for diagnostic-prescriptive emphasis included children's mathematical knowledge and skills, levels of understanding, general interests and attitudes, and learning styles.

Field Experience Treatments

Subjects in both field experience groups were required to plan and conduct eight mathematics sessions and to prepare a project that developed one or more mathematical concepts for instructional purposes.

The subjects in the control field experience group planned and taught eight lessons to two or more children and prepared a teaching unit on a mathematical concept of their choice.

In the experimental group, subjects conducted four diagnostic-prescriptive sessions for a child of high mathematical ability and four for a child of low mathematical ability in the same grade level. Subjects prepared a project that included a log, lesson plans, and materials developed for the sessions.

The Experimental Design

The Design

Four factors were considered in this study: instructional method, field experience method, mathematical achievement, and instructor. Each factor contained two levels and all factors were fully crossed, creating a 2 x 2 x 2 x 2 factorial design. The levels within each factor were:

Instructional Method
  Experimental (ME)
  Control (MC)

Field Experience Method
  Experimental (FE)
  Control (FC)

Mathematics Achievement
  High (AH)
  Low (AL)

Instructor
  #1 (I1)
  #2 (I2)

This study used a mixed-effects model for the design: fixed factors were instructional method, field experience method, and mathematics achievement. Instructor was treated as both a random and fixed factor. The design is shown in Figure 1.

Instructional-Field Experience Combinations

The two field experience methods are fully crossed with the two instructional methods. As a result, there are four separate method combinations:

Instructional method	Mathematics achievement	Instructor	Field experience method	
			FE	FC
ME	AH	I1		
		I2		
	AL	I1		
		I2		
MC	AH	I1		
		I2		
	AL	I1		
		I2		

Figure 1
Experimental Design

1) experimental instruction-experimental field experience (ME FE),
2) experimental instruction-control field experience (ME FC),
3) control instruction-experimental field experience (MC FE), and
4) control instruction-control field experience (MC FC).

Selection and Assignment of Subjects

At the beginning of the semester students enrolled in TED 421 were randomly assigned to the sixteen cells in the experimental design. The assignment process was conducted in three steps. First, each student selected a local elementary school for his or her field experience. At the beginning of the semester, all students with their field experience at nine designated schools were assigned to methods classes in the morning and to the field experience in the afternoon; students who selected the other nine schools were assigned to afternoon methods classes and to the field experience in the morning. After seven weeks the times of day for class and field experience were reversed. Second, test scores from the Arithmetic Test of the Cooperative Mathematics Tests (ETS, 1962) were used to stratify students into two levels of mathematics achievement (high and low). The final step was random assignment to the four combinations of instructional and field experience methods.

Instrumentation

Nine instruments were used to assess the subjects' performance on the following nine dependent variables:

a) knowledge of teaching mathematics,
b) knowledge of diagnostic-prescriptive teaching of mathematics,
c) knowledge of diagnostic teaching of mathematics,
d) knowledge of prescriptive teaching of mathematics,
e) use of diagnostic-prescriptive techniques to teach mathematics,
f) use of diagnostic techniques to teach mathematics,
g) use of prescriptive techniques to teach mathematics,
h) attitude towards mathematics, and
i) attitude towards teaching mathematics.

There are three types of instruments: those which measure knowledge (for variables a through d) those which measure use (for variable e through g), and those which measure attitude (for variables h and i). A tenth instrument was used to determine the subjects' level of mathematics achievement for stratification purposes.

Instruments Measuring Knowledge

At the end of the course all students took an objective multiple-choice final examination. The portion of the examination that covered regular course content was developed to serve as the instrument to evaluate the subjects' knowledge of teaching mathematics, dependent variable a.

Another portion of the final examination was concerned with diagnostic-prescriptive teaching. It included subscales that covered 1) diagnosis (diagnostic content and application of diagnostic techniques), 2) prescription (prescriptive content and application of prescriptive techniques), and 3) the combination of diagnosis and prescription. The three diagnostic and prescriptive instruments were developed by the researcher, who was one of the course instructors, and were used to measure dependent variables b), c) and d). The subscales consisted of written objective multiple-choice items, with one variation. Most of the items in the subscale concerned with application of diagnostic techniques were based on video tape presentations of children in mathematics diagnostic sessions. The subjects observed each video taped session, then read and answered objective multiple-choice questions about the session. Reliability estimates for the four instruments measuring knowledge were obtained using Hoyt's procedure and are presented in Table 1.

TABLE 1

RELIABILITY ESTIMATES
INSTRUMENTS MEASURING KNOWLEDGE

Instrument	n	k^a	r_{xx}	s_e
Regular course content examination	77	27	0.63	1.74
Diagnostic-prescriptive examination	77	31	0.66	2.29
Diagnostic examination	77	19	0.53	1.77
Prescriptive examination	77	12	0.50	1.38

k^a = number of items in the measure

Instruments Measuring Use

Each subject involved in the experimental field experience developed a diagnostic-prescriptive project in which two children, one of high mathematics ability and one of low mathematics ability, were diagnosed, and prescriptive follow-up lessons were planned and conducted. The project was used to evaluate the subject's performance on the three dependent variables, e) through g), pertaining to use of diagnostic-prescriptive techniques.

Two raters, the course instructors, evaluated the projects. Three scores were assigned to each project: one for diagnosis, one for prescription, and the third the combination for diagnosis and prescription. Each rater was blind to the instructional method received by the subjects whose projects she evaluated. Reliability estimates were computed for each of the three ratings. They are summarized in Table 2 where r_{xx_1}

gives the reliability estimates for an individual rater and r_{xx_2} gives the reliability estimates for the combined raters.

TABLE 2

RELIABILITY ESTIMATES
INSTRUMENTS MEASURING USE

Rating	r_{xx_1}	r_{xx_2}
Total Project (Diagnostic and Prescriptive)	0.87	0.93
Diagnosis project	0.95	0.97
Prescription project	0.77	0.87

Instruments Measuring Attitudes

Items to measure dependent variable h), attitude towards mathematics, were selected from existing measures. The items were organized into two subscales: 1) enjoyment of mathematics, and 2) anxiety-confidence regarding mathematics.

Some items to measure dependent variable i), attitude towards teaching mathematics, were selected from existing scales and some items were constructed by the author. The items were organized in four subscales: 1) enjoyment and anxiety-confidence regarding the teaching of mathematics, 2) attitude towards diagnostic-prescriptive teaching, 3) attitude towards teaching for understanding, and 4) attitude towards teaching with manipulative materials. The reliability estimates for the two instruments are summarized in Table 3.

TABLE 3

RELIABILITY ESTIMATES
INSTRUMENTS MEASURING ATTITUDES

Instrument	n	k^a	r_{xx}	s_e
Attitude Towards Mathematics				
Pretest	77	18	0.96	2.95
Posttest	77	18	0.93	3.01
Attitude Towards Teaching Mathematics				
Pretest	77	36	0.81	4.73
Posttest	77	36	0.90	4.60

k^a = number of items in the measure

Instrument Measuring Mathematics Achievement

The Arithmetic Test of the Cooperative Mathematics Tests, Form B (ETS, 1962) was administered to all students prior to their enrollment in TED 421.

Analysis Techniques

All hypotheses tested by this study were non-directional. They were all tested for significance at the α =0.05 level.

Analysis of covariance (ANCOVA) was used to analyze all data gathered on the nine dependent measures and their subscales. Grade point average prior to the course and pretest scores on the attitude towards mathematics measure and on the attitude towards teaching mathematics measure were used as the three covariates to further adjust for initial differences between cell groups.

The ANCOVA mixed-effects model was used for data analysis. The design was simplified by pooling, however, when interactions with instructor were considered not significant.

A Dunnett test was used to compare the four instructional method-field experience method combinations. These multiple comparisons were made for all dependent variables with a significant main effect for instructional method or for field experience method, or a significant instructional method by field experience interaction. The control instruction-control field experience method combination, MC FC, was compared to each of the other groups, ME FE, ME FC, and MC FE.

Summary of Results

Effects of Instructional Methods

The most interesting effects of the instructional methods concern diagnostic-prescriptive techniques for teaching mathematics. Four findings indicate that instruction in diagnostic-prescriptive teaching does significantly and positively affect the preservice teacher's knowledge and use of techniques for teaching to the needs of individual children. The first three findings are on the following measures: the diagnostic-prescriptive examination (used to evaluate overall knowledge of diagnostic-prescriptive teaching), the diagnostic examination (used to evaluate knowledge of diagnostic teaching), and the diagnostic project (used to evaluate use of diagnostic techniques). The analyses of the diagnostic-prescriptive examination subscales indicate a significant difference on the composite scale for application of diagnostic-prescriptive techniques. It should be noted that no significant difference was identified on any measure that deals exclusively with prescriptive teaching.

There were also no significant differences on attitudes or on knowledge of teaching mathematics, measured by the regular course content examination.

Effects of Field Experience Methods

There were no significant differences due to field experience methods on the tests of attitude or on the four main tests of knowledge. The subscale measuring the application of diagnostic techniques did, however, produce a significant difference favoring the experimental group. This result indicates that using diagnostic-prescriptive techniques with individual children in actual teaching situations does give preservice teachers valuable diagnostic experience which readily transfers to other situations where application of diagnostic techniques is required.

Effects of Instructional-Field Experience Method Combinations

The preceding analyses of instructional methods and field experience methods demonstrate that preservice teachers' skills in diagnostic teaching are significantly influenced by the experimental methods of both instruction and field experience. The findings, however, address each experimental method independently of the other experimental method. Detailed analysis of the various combinations of instructional and field experience methods revealed that significant differences regarding diagnostic skills are caused by a specific combination of methods.

Multiple comparisons among the four methods combination groups were made on the diagnostic-prescriptive examination, the diagnostic examination, and several subscales. The experimental instruction-experimental field experience (ME FE), experimental instruction-control field experience (ME FC), and control instruction-experimental field experience (MC FE) groups were each compared to the control instruction-control field experience (MC FC) group. Both mixed methods combinations (ME FC and MC FE) failed to exceed at a significant level the control methods combination (MC FC) on any of the multiple comparison tests. But when experimental instruction is coupled with experimental field experience (ME FE) the combined treatment is superior to the combined control treatment (MC FC) for knowledge of diagnostic teaching ($p<0.05$).

Effects of Mathematics Achievement

Findings on the diagnostic-prescriptive examination and project indicate that the level of mathematics achievement influences the preservice teachers' knowledge, but not use of techniques for diagnostic-prescriptive teaching. Specifically, results on the examination and its subscales reveal that subjects in the high mathematics achievement level performed significantly higher than those in the low achievement level on the total diagnostic-prescriptive examination, the diagnostic examination, the diagnostic content subscale, and combined diagnostic-prescriptive content. There was no significant difference on any of the measures that deal exclusively with prescriptive teaching.

There are several explanations for the differing effects that mathematics achievement had on diagnostic and prescriptive knowledge outcomes. First, techniques, for diagnosing mathematical strengths and weaknesses require a firm understanding of the mathematical concepts involved. Unlike the usual methods approach which emphasizes "how to teach", diagnosis also emphasizes "what to teach". A preservice teacher who has mathematics deficiencies may find it difficult to understand the use of diagnostic techniques because of the mathematics involved in determining the "what to teach". In simple terms, in order to know that something is wrong, the diagnostician must know what is right. Similarly, diagnoses of higher ability children may confront the teacher with situations that require more than a modicum of competence in mathematics. The significant findings on the diagnostic knowledge measure bears out this line of reasoning.

Interactions Among Factors

There were no significant interactions on the four knowledge instruments.

Results concerning interactions on the projects which measured use of diagnostic and prescriptive techniques are mixed. There is no significant interaction on the diagnostic project. But, a disordinal instructional method by instructor interaction is significant on both the prescriptive portion of the project and the total project.

There are significant interactions on the attitudinal measures. On the attitude towards mathematics scale, these include the instructional method by field experience method interaction, the instructional method by mathematics achievement interaction, and the three-way field experience by mathematics achievement by instructor interaction. Of further interest is the fact that these three interactions are also significant for the measure's anxiety-confidence regarding mathematics subscale.

Subjects in both the ME FE and MC FC method combination groups scored higher on mathematics attitude that did those in the mixed combination groups, ME FC and MC FE. One explanation for this instructional method by field experience method interaction is the following. When instructional and field methods complement each other, preservice teachers feel that they are practicing what they are studying. This is conducive to positive feelings about their learning experience and about their topic of study. Such learning is pedagogically conducive to producing positive attitudes towards the topic of interest -- in this case, mathematics. But when instructional and field methods do not complement each other, as is the case for ME FC and MC FE, preservice teachers could feel frustrated and anxious about mathematics.

Conclusions

The findings of this study indicate that both instruction and field experience in diagnostic-prescriptive teaching of mathematics are viable as components in a preservice mathematics methods course. This conclusion is based on the following outcomes of the study.

1. Instruction in diagnostic-prescriptive methods can be taught successfully to preservice teachers. Evidence supporting this statement comes from four significant differences that favor the diagnostic-prescriptive approach over the control approach. First, and foremost, the experimental instructional subjects exceeded the control instructional subjects on the overall test of diagnostic-prescriptive teaching of mathematics. Second, scores of experimental subjects were superior to scores of control subjects on their knowledge of how to apply diagnostic-prescriptive techniques. Third, scores of experimental subjects were superior to scores of control subjects on knowledge of diagnosis, which covers both content and application of diagnostic techniques. Finally, on the diagnostic-prescriptive field experience projects, scores of experimental subjects surpassed scores of control subjects in using diagnostic techniques.

2. Field assignments in diagnostic-prescriptive teaching provide preservice teachers with valuable experience for learning how to diagnose children's mathematical strengths and weaknesses. This statement is supported by the results on the application of diagnostic techniques subscale, on which the experimental field method subjects performed significantly higher than the control subjects.

3. Field experience using diagnostic-prescriptive techniques provides positive reinforcement of diagnostic-prescriptive instruction. Three points support this contenton. First, the refined analysis available through the multiple comparisons test reveals that the experimental instructional subjects need the corresponding diagnostic-prescriptive field experience in order to maintain their superiority over the MC FC group on the dianostic examination. The ME FE combination is superior to the control combination, while the ME FC combination, which lacks the reinforcing field experience, is not. Second, the test of the subjects' attitude towards mathematics indicates the possibility that the subjects' attitude is more positive when the field experience and instructional approaches are matched (ie., ME FE) than if they are mixed. Finally, from a pedagogical perspective, the linking of experience to subject matter makes sense. Learning and applying a unified approach to teaching provides consistent reinforcement of the approach in both the university classroom and the elementary schoolroom.

4. The diagnostic-prescriptive approach can be taught without sacrificing the preservice teachers' understanding of more customary mathematics methods. The absence of a significant difference between experimental and control instructional groups

on the test of regular course content supports this conclusion.

Extensions of the Experiment

Research on instruction in diagnostic-prescriptive techniques of teaching mathematics with preservice teachers is limited. In the course of conducting this study, several ideas developed for extending its scope.

One of the more obvious extensions of this study is a follow-up study to determine the long range effects of training preservice teachers in diagnostic-prescriptive teaching of mathematics. Follow-up of preservice teachers could be conducted during the student teaching experience or later in the teachers' careers. Such a study would determine if they are using the diagnostic-prescriptive approach with children in their own classrooms. Although difficult to conduct, such a study would provide valuable information regarding the usefulness of diagnostic-prescriptive instruction for preservice teachers.

Field experience provides practical learning situations different from those of the college classroom. But, it is difficult to provide teaching models which utilize the instructional emphasis of the mathematics methods course. An interesting question to investigate regarding this problem might be the effects of assigning preservice teachers who receive instruction in diagnostic-prescriptive teaching to two groups of teachers in the field experience: one group that models a diagnostic-prescriptive approach to teaching, and one that does not. A variation of this idea is to couple instruction in diagnostic-prescriptive teaching with a field experience in a mathematics clinic for elementary school children.

The attitude measures used in this study detected no differences in subjects' attitudes towards mathematics or attitudes towards teaching mathematics. A third attitudinal variable that might be used in future research is preservice teachers' attitude towards children. Diagnostic-prescriptive teaching encourages close teacher involvement with school children, and that may affect teachers' attitudes towards children.

Perspective

The findings of this study, and other studies on related subjects, are building a base for a change in preservice education. Many diagnostic-prescriptive techniques are effective in meeting the needs of children, and such techniques are teachable to those who will instruct children. Perhaps the time is coming for teacher educators to develop a more systematic approach to meeting the learner's needs through the training of preservice teachers in diagnostic-prescriptive teaching.

Diagnosis and Prescription
for the
Research Council on Diagnostic and Prescriptive Mathematics

Gerald R. Rising
State University of New York at Buffalo

dedicated to John W. Wilson

I am not a member of the Research Council on Diagnostic and Prescriptive Mathematics. On the other hand I am not a disinterested observer of your activities. I believe that yours can be the most important mathematics education organization. In what follows I will be attempting to push you in that direction.

As I address this task of analyzing your activities I believe that I can provide a useful service to you by raising some serious questions if you are willing to consider them now before it is too late.

That adverb now is most important. You are not yet, I hope, completely fixed in your organizational direction even though I see signs that you are moving far too rapidly toward that rigidity that first stultifies and eventually stymies progress. It is my sense of the situation that despite this rapid growth you all cite so regularly you are in fact careening along toward your imminent destruction. This has led me to take my task most seriously. I urge with all the emphasis I can place on the request that you consider this matter equally seriously.

As homework in preparation for this analysis I have read with care as many of your documents as I could lay my hands on. This study has included your constitution, all issues of your newsletter, four collections of conference papers, and two D/P texts.

The Quality of Your Writing

I find myself quite frankly aghast at the low quality of your publications. With a very few exceptions they present an extremely poor image for the Research Council. Nothing is to be gained by a claim that these papers are not meant to be finished products. That would miss the point for whether you want them to or not these papers do represent you. Too much of the writing lacks the level of basic literacy that we should expect in a thank you note from a high school student. How can you, many of you university professors of high rank, allow your names to be attached to these papers? I am afraid that I know the answer to that question and I will return to it at the end of this paper but now I will explain and document my case.

Start with the basics. The editing is terrible: an average of worse than one typo per page, misspellings everywhere -- it took three volumes for consensus to lose a second "c", and the simplest of words misused: for example "sight" for a location. There is no evidence of recourse to a dictionary. Lock-jaw sentences abound even though, as I understand it, these papers were all presented orally. In several papers paragraphs are assigned at random. In the worst of these the generator appears to be set at two sentences per paragraph. I looked twice to see if this was a textbook author trained to love white space.

Now those bloopers would not be so bad standing by themselves, but they appear in papers in which the language is preposterously pedantic. Consider only a few examples. "Wherein subjects who learned a superordinate task were found to no longer retain some supraordinate tasks which were known to be prerequisite." When we ought to be clarifying, instead we give every appearance of trying to baffle our audience. I envision attentive members of an audience struggling with those words and missing a paragraph or two of what follows. It is not a satisfactory response to claim that there is a restatement in the next sentence: the first should then never have appeared.

Another example of this purposeless confusing of the listener is "However, there now exists an abundance of research which refutes the objection that application of ability training techniques does not increase achievement." You can make sense of it on rereading -- listening the first time, I doubt it. The next sentence is upon you too soon.

But much more of this effect is due to overuse of technical and pseudotechnical jargon. One example of this made me especially angry. One writer uses the word "praxiologists" (sic). This led me on a futile search through all my dictionaries. Unsuccessful, I set out again, this time checking every word beginning with "prax". Still not in the first three which included the Oxford English Dictionary, I found it finally in the Third New International with an "e" for that first "i". Now after all that to the meaning. It means people who study events or what happens. I couldn't believe it: good God, don't we all?

There is a very interesting transition in one paper from this same kind of prose pomposity: antecedent, stimulus resources, descriptor, consequent (as a noun), and discrepant -- the clause is "if the response is judged to be discrepant" which translates for us common folk to "if he doesn't know" -- in the writer's discussion of material with which he feels very confident, to a much clearer and attractive vocabulary when he turns to a newer area of investigation. Here he talks of planning, memory, coding, perception. I believe that there is an important message in this for all of us. As he reviews material in which he has less interest the writer is fluffing his feathers with unattractive jargon, but when he wants to convince us and so is more concerned about communication he addresses the content in direct clear English. I urge all of you to adopt the second mode. The easiest way to do so is to take yourselves out of the lecturer role and into the learner role.

Of course there are exceptions to this. One of the best of these is a very useful and suggestive paper on linguistics. I take the liberty of displaying once again what bothers me by translating one of its clear statements into the language of RCDPM. The writer says: "...in both cases the child must bring meaning to the words spoken -- to the natural language and to the words that are specifically mathematical terms." RCDPM would have had that: "under alternate circumstances the pre-operational learner must transform into cognitively manipulable and applicable structures the learning units -- to both vulgate and particularistic content associated descriptors."

Now that is the first level problem; it is not unimportant, however, merely because it is tied to what some of you might consider prosaic matters. In fact it is especially important in this organization which brings together practitioners -- clinicians and classroom teachers -- and researchers. It is, I believe, exactly the kind of thing that I have been talking about that leads Martha Hopkins to call for a "training session for non-researchers presenting... various techniques and terms used in research (and its interpretation.) This would provide the general classroom teacher with the opportunity to understand, interpret, and evaluate the research studies...." Such a session should never be necessary. In fact I propose an alternative: a presession staffed by one of our English graduate students who would repeat for the researchers of this organization the remedial composition course for freshman general education students he teaches on campus. At the very least I urge each member of this organization to purchase or borrow and study one of those pocket guides to clear communication, like the one by Rudolph Fleisch. Then those who write can apply the lessons and those who read can better judge the product.

I move next to equally serious problems that relate to the ideas of your authors. Here again I find two levels of writing. Whenever you talk about specific activities with students you are clear and usually directly to the point. There are never enough of these straightforward descriptions. In fact they seem to embarrass you as being too informal, not researchy enough. That I find sad, because here is where you are at your very best.

It is when you turn to analysis and conceptualization that you are extremely weak. This is, I believe, an artifact of our recent history in training graduate students in education within the frameworks of behaviorism and the controlled experiment. Examine any publication that comes from that tradition -- and there are too many here -- and you will find:

(1) Endless and mindless reviews of "the literature". Even though we know that most of the studies cited are worthless or in Romberg's less severe terms "poor but improving", these are quoted as if tablets from Sinai.

> The only time an article is criticized is when it is juxtaposed with a paper which offers contradictory results.

Then after the data presentation:

(2) A results section that at best repeats data.

And finally:

(3) A conclusion and discussion section of a paragraph or two which offers virtually no help in placing these results into a larger context.

What is so clearly and purposefully preened from such documents is careful analysis by the author. My estimate of the extent of author analysis in current educational research literature is less than five per cent. For those few social scientists who still believe, despite thousands of tons of evidence to the contrary, that the experiment itself must and will tell all, that approach is justified. For the rest of us it is not. We must begin to think again.

I believe that in these papers authors are indeed thinking once again and I credit them for that. I guess we should expect our early steps in this direction to be hesitant and unbalanced. For that reason I offer the following criticisms not at all in the same condemning spirit of my earlier comments but rather in the clear hope of raising the level of dialogue.

These papers abound in what I designate pedagogical bromides (my colleague Steve Brown calls them pedagogical truisms): trite statements too easy to accept -- and often generally accepted -- but for which at least as much can be said against as in favor.

Consider the assumptions that one writer claims are necessary to identification of D/P research:

(1) Strengths and weaknesses can be identified as an evaluation of a child's position on a continuum of skill development (Gagné).

(2) These... can be assessed....

(3) These strengths and weaknesses are generally conceptual and (relevant) to instruction."

The first of these statements is patently false as others (Underhill for example) have shown. The writer assumes uniqueness and linearity here when such assumptions are unwarranted. The second is demonstrably false because of the difficulty with (1). And the third is only acceptable in such a general way that it is not useful.

A trio of authors write "Ideally, all teaching should be individualized because learning is an individual matter." The insistence on universals makes this clearly false, but even if we loosen "all" to "most" I'm not satisfied. How does a student learn to defend his arguments against peers if he is allowed no peers? Interaction and yes, even competition are useful supports to good learning and teaching.

Or another writer, "It is the thesis of this paper that understanding mathematical notions can be expressed through actions, images, and symbols and that these alternate expressions of understanding may be conceptualized as levels of maturity." That sentence is typical of the bromide. If you relax at all it drifts by you, but if you try to capture its essence, like mercury, it squeezes away.

In experiments too we can produce pedagogical bromides. Thus two of your experimenters prove once again that something is better than nothing; and another proves that something related to criterion is better than something unrelated to criterion -- these in the 1970's!

It should be clear that one of the answers to the problem of pedagogical bromides is the same as the answer to our first series of problems: tighten up the use of the English language. Greater directness and clarity will usually expose where these arguments lose their force. At the same time such tightening will help us to get at the more fundamental concepts and issues to which we will turn next.

But there are other tests that we should learn to apply to our pedagogical thinking in order to avoid these superficially attractive but basically misleading ideas. Here are just a few:

(1) Imagine a setting in which your wish has come true. Is the result necessarily better?

(2) Think of situations that run counter to yours. Are they all necessarily bad?

(3) Try to fit some extreme cases into your setting. Are there alternatives that are better for these cases?

As I write even these few suggestions I worry about their effect. I offer them to encourage you to think more deeply and in new ways. I do not offer them to encourage you to hedge every statement with lists of "except fors" and "only ifs." This latter road leads back to empirical thinking where the value to educational practice declines sharply as the controls fall into place.

Conceptual Problems Related to Diagnostic and Prescriptive Mathematics

I have dedicated this paper to John Wilson through respect gained over the years for this friend. That respect was significantly -- but sadly -- enhanced when only after his death I read his epistomological paper in support of the clinical intervention research strategy (P3, 1-24). His essay is a model of clarity and stands as a nearly perfect answer to my criticisms of writing in the last section. Here is tough content made easy by a genuinely literate teacher.

We will miss this fine colleague. I am quite frankly surprised to have found no reference to the low-stress algorithms John and others used in any of the D/P papers or books I examined. I wonder how soon someone will rediscover that remedial wheel.

But I cite John here for a quite different reason. He has provided each of you who uses clinical intervention as a research strategy with all the justification you will ever need. His seminal paper should put to rest your feelings of inferiority that have kept bubbling up in your apologies for this important tool. John's paper makes it abundantly clear that this method is if anything more valuable than the controlled experiment. The fact that unlike the experiment it occasionally provides useful and applicable information for teaching should certainly not count against it.

Do not take these comments as complete agreement with the cast of John's argument for I would go further and would opt for emphasis on still other forms of research, but clinical intervention is exactly right for your work and much of my own.

However, I believe that it is very important to point out that this research technique is not congruent to diagnosis and prescription. My sense of the difference is that they cut orthogonally to each other like lines of latitude and longitude. A clinician, for example, would diagnose problems for and provide remedial assistance to many patients without any recourse to clinical intervention research. But after a time he -- and quite possible others -- might notice some common feature of his work for which he could evolve and refine and test one or more intervention strategies. Now as he accumulates evidence he is applying the clinical intervention. In both there is much of what educators today are calling the wisdom of the practitioner but in clinical intervention that wisdom is refined and sometimes even changed.

Clinical intervention research is like much of medical research. (I note here that they and we will still continue to use experiments when possible and appropriate.) Diagnosis and prescription is also an honored technique in medicine. That is not unexpected since education has taken the terms from the M.D. But the analogy is not complete The doctor's role is to restore the patient to health and usually has little to do with growth. The educator on the other hand is very much concerned with growth. Underhill (I, 127-130) tries to retain the analogy by making the associations:

etiology ⟷ etiology

disease ⟷ problem

symptom ⟷ error

where by etiology we mean the search for causes. I do not believe that this suffices because -- to take an extreme case -- you could straighten out all the identifiable problems, that is error patterns, and still leave a child who will not function well as he meets new mathematics. At the opposite extreme you can have a student who appears to make little or no progress at the error or problem level yet is making great gains in growth .

The medical analogy breakdown is related to our even more central problem, that of the persistent and pernicious belief that the end of the rainbow is in a pot of hierarchies. Certainly no one would claim that there is no logical order to mathematical pedagogy at all. I would indeed not want to teach the solution of general quadratics before counting. But the hierarchy model has captured too much of you completely. Here are some of the simple things wrong with this model:

(1) There is no one best route, either linear or latticed, through mathematics.

(2) Mathematics learning is not a binary matter -- all or nothing, learned or not learned -- even when considered in small pieces.

(3) Under the best of conditions tracing individual problems back down even local hierarchical paths proves a miserable failure as a remediation device. When Mary Robinson was still Roy Callahan's excellent student, she demonstrated to us how illusory was this hope.

For these reasons and two others I recommend that you move away from these hierarchy constructions. My first reason is that you would be well advised as a remedial teacher to respond to the order of the student's classroom text which after all is the sequence, at least beyond your grade, to which the student will return.

Here I had better make clear my position regarding teachers and textbooks. I believe that a classroom teacher has a particular kind of responsibility in the scheme of education, as a matter of fact the most important kind of responsibility. The classroom teacher has plenty to do and should not be expected to be either a curriculum developer or a clinician. Instead the teacher should be provided with an instructional program, usually in the form of textbooks and associated learning materials, and given the responsibility of implementing that program carefully and thoughtfully. Where I see departures from this approach I generally see disasters in the making and here among the departures I include locally developed individualized programs, levels programs, computer based resource units, locally developed curricula (which usually have no effect on program other than wasting teacher time and energy), and programs organized around diagnosis and prescription.

Tinkerers with textbook programs in school district and university offices ought to be required to sit down with textbook authors like Vincent Glennon or Tom Romberg or Lelon Capps to see just how much effort is devoted to the generation and implementation of a logical developmental sequence for mathematics that covers the full span of grades. Heaven knows that textbooks are not perfect, but I have yet to see anything else one-fifth as good.

A colleague at Buffalo State College, Russ Macaluso, took over direction of their laboratory school only to find himself saddled with a staff whose whole teaching program was organized around the D/P approach and with school achievement scores which he felt were far below what should be expected for this school. Despite some early staff resistance and the expected transition problems, he moved them quickly into a more conservative program and has improved his achievement record in the process. He reports to me that the school is in better order; the teachers, students, and parents are happier; and even the most reluctant teachers admit that they are accomplishing more with less energy expenditure. This last is especially important: I see teachers in some of our area schools that are involved in these wildcat programs who are literally killing themselves. At the same time their students are receiving negative pay-off, all because some principal or supervisor has taken a curriculum course and wants to show off his pet -- of this year.

Now let me return from that tirade to the matter at hand, hierarchies. I have given my first reason beyond the failure of the model for turning away from their construction, to wit preference for the order of the program of which the student is already a part. My second reason relates more closely to the failure of the model. There is such a thing as usng too high a power on your microscope. That is exactly what I believe Hill and Underhill do when they refine addition beyond the point of even clinical usefulness. I hope that clinicians will respond to this point. Do you not find that problems at the detail level are more often a matter of hazy more general notions preventing the students from pinning down specifics rather than specific atomic matters like all the uneven divisions in the tens or hundreds place?

This is the first practical problem. There is also a second. Diagnosis, I submit, is not an activity that should stand alone. This is in fact one of my great difficulties when for example I read that a writer was recording all unusual student responses. What are these to be used for if not remediation? Are they to serve like those picture postcards of national parks for entertainment after retirement? No. What diagnosis should serve are: (1) development of a remedial response for that student, and (2) over the longer haul development of more general remedial or even original teaching regimens. The first of those is the D/P response, the second the clinical intervention response.

Several remarks in my discussion will have led many of you to believe that I am totally opposed to diagnosis and prescription. That would be a misunderstanding of my position and I will attempt to clarify my posture now. I believe with Callahan that prevention is far more important than remediation. But when youngsters need remedial assistance we should provide it in the best way possible. For remediation then, I support you in seeking the most rational avenues to response. I will criticize you as I have already and as I will soon again for some of those avenues, but I believe in your purpose and your process. For original instruction, on the other hand, application of D/P techniques constitutes a hoax. You can get away with it only because so much in the school mathematics curriculum is repetition. The imposition of D/P on to original instruction is at once illogical and very low or negative in cost-benefit return.

I have no final answers to give on the question of the clinician vs the classroom teacher. I worry wherever I see two neighboring classrooms, in one of which a teacher struggles with thirty or more students while in the other another teacher works with one student. Would not our benefit/cost ratio be improved if we split the class in two?

I now turn to what I believe is the tremendous imbalance in your activities between diagnosis and prescription. When I look at your work to date I can see no justification for the fifth letter in your acronym. It should be simply RCDM. I believe that that constitutes a very serious failing on your part. If you continue on your current course you'll end up with clinics filled with computer banks and statisticians and no teachers! Why has this imbalance come about? Here are some reasons:

(1) Diagnosis is far easier than teaching.

(2) Diagnosis as currently practiced applies those educational psychology skills that have had such a position of honor in the recent past.

(3) Diagnosis involves no commitment, no possibility of failure on the clinician's part. That comes only when a response is offered.

(4) Diagnosis has been associated with content ordering that is familiar to everyone. Following the tradition of mathematics methods texts of fleeing from pedagogy to content, so too does the diagnostician.

(5) Diagnosis as generally perceived is an essentially mechanical process.

I challenge you in this organization to bring your activities more than back into balance. I urge you to devote at least two-thirds of your time to the prescription-remediation end of this work.

Where are the case study series that we ought to be generating? Where are the reports of successes and failures with individual students? Where are the follow-up studies? Where are the detailed recommendations for what to do? If there are any of these, you're keeping them well hidden.

The general approach to prescription is simply reteach. A typical generalization from one book is: "Good remedial teaching or corrective teaching is similar to good initial teaching in that it adheres to basic principles of learning." Another bromide: essentially meaningless. Other suggestions follow: "Mastery of the basic facts requires practice. The facts must be studied and memorized. The student should not develop the habits of counting on fingers, tapping, or relying on the number line. If certain facts are troublesome, he or she should concentrate on those. Student progress should be measured by checkups...." After a bookful of diagnosis, that's not much help. The one useful suggestion offered later is the think aloud technique. Reisman's book (2) is much better but still must leave the neophyte uneasy about any but the particular lines of her examples.

What remediation is undertaken focuses tightly on specific errors -- current diagnostic techniques after all lead to this. And exactly here is where the symptoms-instead-of-the-disease problem is introduced. The disease is often much more than Speer's mosaic of errors, it rests in deep psychological, sociological, or neurological problems. Reteaching how to borrow will just not do in such cases.

Over the past several years I have had the fortunate opportunity to observe Frederique Papy many times working remedially with small groups of students. Because her work is very much to the point here, I will describe it to you briefly. She met her two groups of four students each for about thirty minutes on alternate days. The children were boys and girls from central St. Louis, by far the weakest in their primary classes. Frederique's basic tools in working with these children were hand held calculators (used mostly for record keeping and as checking devices) and the other pedagogical devices of the CSMP elementary mathematics program: strings, arrows, and the Papy minicomputer.

I digress briefly to comment on the CSMP elementary program only as it applies here. This program is non-linear. Students cycle through activities many times until they can build the algorithms of mathematics on extended experience. Whether this program is successful or not you must ask yourselves how does your linear model for diagnosis fit it? I suggest that is does not, which leads us quite appropriately back to Frederique's remedial classroom.

I have worked in ghetto schools before and have seen many very weak students but none as bad and as seemingly unremediable as these. When I first observed these second graders Frederique had them playing what to most students would have been a trivial game. At the time of my first visit the students had been playing this game for several weeks but were still making very tentative moves.

The dynamics of the groups were very interesting. Whenever Frederique asked a question, all four hands would fly up and the children would rise up out of their chairs. But as soon as she called on a student that student would recoil into himself giving almost the physical appearance of a snail retreating into its shell. Only after great hesitation would the student offer a tentative response.

Now break with me for four months as I did until my next visit. Before attending the remedial class I was shocked to see one of these same children in a regular class participating voluntarily. In the remedial groups the students were now seriously and reasonably addressing the kinds of activities -- different by now of course -- that Frederique offered them.

What had happened? I have talked about this at length with Frederique and other CSMP staff members. Frederique believes, and I agree fully, that many of these children have severe emotional blocks that are preventing them from even addressing mathematical problems. A good share of them derive from home settings, for example, which will not go away and to which the school cannot respond. We can only overcome those blocks through long and patient instruction. The students need to be made to feel secure through steady even-handed treatment and when they are successful they should receive encouragement. They need to be set tasks which they can fulfill, but at the same time they must be led to understand that they must live up to expectations. The competition used as a motivating device must be controlled so that it does not also act as a punishing force. As they work the students must gain trust in mathematics as solid and not shifting ground, something that will work for them.

Frederique told me that she could identify a plateau effect among these students. They would seem to make no progress but suddenly an individual student would break through his barrier and from then on make rapid progress. For some this took only a matter of days, others weeks, and one boy many months. The trigger for this discontinuous leap appeared to be tied to time as uniquely identified for each child.

You will notice, of course, that Frederique's remediation regimen fits nicely with the CSMP program itself. Both operate on the basis of cumulative effect. I do not believe that the association is a necessary one and I urge you to experiment with her approach to remediation especially with students with severe learning problems. Her activities and a few comments about students are included in the CSMP books: Math Play Therapy (3). I urge all of you to purchase, study, and experiment with the regimens she introduces in those books.

It appears clear that we have something to learn also from the neuroscientists, the brain folks. While I have reservations about many of the earlier predictions of breakthroughs in the brain revolution, I do believe that there is some interesting work being done here that relates to your enterprise.

Fundamental Organizational Issues

I have only a few comments to make here and I will save my most serious concern for last.

Shortly after I was appointed to the NCTM Research Advisory Committee, a friend of mine caught me in a hallway at an annual meeting. He congratulated me briefly and then moved directly to his real concern: he urged me to use my influence to increase the number of pages annually in the Journal of Research in Mathematics Education "Why?" I asked him. "Look," he responded, "you know what we're up against. We've got to get articles published to survive." Some of you may recall that I asked earlier why some university professors were willing to associate their names with the low quality papers they were turning out. There I am afraid is the answer: quality doesn't count for much today, only quantity.

Despite that and in an odd way because of that, I recommend that this organization sponsor a journal. With tight quality control a journal can provide you with an opportunity to share your best work. More important it can provide you with an opportunity to redirect attention in mathematics education toward problems and methods of responding to those problems which are more attuned to your and my interests than most journals currently available. Your situation is parallel to but not quite as bad as that of mathematics educators whose central concern is with teacher education. There is no journal in our field which will carry articles on this topic except with great reluctance. That is one reason why we have little or no active research or development in mathematics teacher education in this hemisphere.

So far, your publications have, I believe, taken the wrong form. You invite someone to talk and promise him a place in a proceedngs volume before you know what is to be said. I urge you to stop that and to turn instead to sponsorship of a refereed journal with a high quality staff.

Callahan has asked "Whether goest RCDPM?" If, as I believe he meant, "Whither goest RCDPM?" for he answers "NCTM". In any case that slip is invested with meaning: I believe that whither you go will determine for you whether you survive.

I disagree with that recommendation about NCTM for a number of reasons, the most important of which is the fact that it could and would create problems for you. You would, for example, immediately find yourselves in conflict for very limited resources with the Research Advisory Committee. As a former NCTM board member, I identify that avenue as a dead end for your organization.

I believe that your organization is needed and should be supported exactly because it can serve a group that is currently not well served by other organizations. What bothers me is that too many of you view your group too narrowly. This is the view that is suggested by the "DP" in your title. It is broadened in the stated purpose of your constitution in a way that I find satisfactory but not yet satisfying. Why? Because I want your organization to catch up in its net all of the currently disenfranchised mathematics educators and not just those who are willing to subordinate themselves to what is your current more specialized interest.

Here then are the constitutional changes that would set you on a better course.

ARTICLE I. NAME The name of this non-profit association shall be International Council for Mathematics Education Research and Development.

ARTICLE II. PURPOSE The purpose of the International Council for Mathematics Education Research and Development is the stimulation, generation, coordination, and dissemination of research and develoment efforts related to the enhancement of mathematics teaching and learning and the prevention or correction of mathematics learning difficulties.

Beyond minor recasting there are only three chages in what I have suggested. The first of course is the organization's title which replaces "DP" by "R" and "D" and adds an "I". The second is the addition of the phrase "the enhancement of mathematics teaching and learning." The third is the removal of reference to the specific "diagnostic and prescriptive techniques." There is precedent for such a change. In England recently, for example, the Association for Teaching Aids in Mathematics became the Association of Teachers of Mathematics.

I do not believe that the minor modifications that I have recommended would change your membership roster quickly, but it would make room for others. I would hope, in particular, that if you follow this first suggestion you would then establish a special interest commission to seek to bring in mathematics teacher educators, mathematics curriculum developers, and others who currently have no comfortable home. This, I am convinced, would not dilute your effort but rather strengthen it.

My recommendations for change would not go far toward solving what I believe is another serious problem of this organization, and that is keeping an honored place for the practitioner in your ranks. I expect, sadly, that you are slowly but surely driving these important members away and I urge you to seek the best ways to encourage them to stay with you. Without them you will find yourselves too closed a group for your own good. I wish I had answers here -- a negative dues asssment perhaps -- but I have nothing to offer.

I have diagnosed. I have prescribed. I will now follow with interest your remediation program, if, that is, I have been able to convince you that any remediation is indeed in order. That is a matter of motivation and every educator knows you don't motivate anyone by criticism. We'll see. For whatever course you choose, you have my interest and my best wishes.

References

1. Heddens, James W. And Aquila, Frank D. Proceedings of the Third National Conference on Remedial Mathematics, 1976, 151 pages.

2. Reisman, Frederika K. A Guide to the Diagnostic Teaching of Arithmetic. Columbus, Ohio, Charles E. Merrill, 1972. 152 pages.

3. Papy, Frederique. Math Play Therapy. Available from CSMP, 3120 59th St., St. Louis, Missouri. 63139.

Editors' Note: In Professor Rising's talk he referred to specific writers and to specific publications as examples of poor writing. We believe the points he makes can stand without these references. It should be pointed out also that no speaker was promised that his or her presentation would be included in this volume.

PUBLICATIONS AVAILABLE

1. The Kent State University Checklist of Mathematical Concepts K-8 (Kent State Mathematics Education Team)
   A 30-page comprehensive checklist of concepts sequentially ordered within broad "concept clusters" such as place value; sets; operations on wholes, integers, rationals; geometry; probability; etc. Newly Revised. $2.00 ($1.60 - members).
2. A Bibliography Related to the Nature of Diagnosis and Remediation in Mathematics (Speer)
   A 67-page, non-annotated, bibliography of recent publications (including research studies) dealing with diagnostic and prescriptive concerns. Entries are arranged by author and are not indexed by content. $2.00 ($1.60 - members).
3. Selective Definitions Relative to Diagnosis and Remediation (Speer)
   A 21-page collection of terms defined in a dictionary format. (New printing of a previously issued monograph) $1.00 (.80 - members).
4. Remedial Mathematics: Diagnostic and Prescriptive Approaches (Higgins & Heddens, Eds.)
   A complete collection of the papers presented at the first National Conference on Remedial Mathematics. May only be ordered by writing ERIC/SMEAC, Room 310, 1200 Chambers Road, Columbus, OH 43212 and requesting item #251M $3.50 — postage.
5. Proceedings of the Third National Conference on Remedial Mathematics (Heddens & Aquila, Eds.)
   A complete collection of papers presented at the Third Conference including a summary of the accompanying invitational research conference. $6.75 (5.50 - members).
6. Clinical Investigations in Mathematics Education (Speer, Ed.)
   A complete collection of the invitational research papers presented at the Fourth Conference. $4.21 ($3.37 - members).
7. Issues for Consideration by Mathematics Educators (Denmark, Ed.)
   A collection of theme papers from the Fourth and Fifth National Conferences with particular emphasis on research of primary interest to mathematics educators. $4.00 ($3.20 - members).
8. An Annotated Bibliography of Periodical Articles Relating to the Diagnostic and Prescriptive Instruction of Mathematics (M. Hynes, Ed.)
   Prepared as an extension of the original RCDPM bibliography, this 46-page publication focuses on recent articles. $3.50 ($2.80 - members).
9. Selected Topics for Diagnostic and Prescriptive Instruction of Mathematics for Classroom Teachers (M.E. Hynes, Ed.)
   A collection of papers from recent conferences with particular emphasis on topics of primary interest to classroom teachers $6.50 ($5.20 - members).
10. The RCDPM 1980 Research Monograph (Romberg, Ed.).
    A 281-page collection of papers from the Fifth and Sixth National Conferences. $11.25 ($9.00 - members).
11. 1980 John W. Wilson Memorial Lecture. (Neuropsychology & the Instructional Psy. of Math.)
    A complete text of the first Wilson Lecture delivered by Vincent Glennon at The Seventh Annual RCDPM Conference. $2.00 ($1.60 - members).
12. 1980 John W. Wilson Memorial Lecture Audio Tape
    $6.75 ($5.50 - members).
13. The 1981 John W. Wilson Memorial Lecture: Diagnosis and Prescription: The Affective Half by William C. Morse.
14. An Annotated Bibliography Relating to the Diagnostic and Prescriptive Instruction of Mathematics, Volume 2 (B. Sadowski, Ed.)

Inquiries concerning any of the mentioned publications should be addressed:

Dr. James Heddens
Publications Sales Manager
441 Beryl Drive
Kent, Ohio 44240